

City University of New York (CUNY)

## CUNY Academic Works

---

Dissertations, Theses, and Capstone Projects

CUNY Graduate Center

---

2007

### The Qualitative Character of Spatial Perception

Douglas Meehan

*Graduate Center, City University of New York*

[How does access to this work benefit you? Let us know!](#)

More information about this work at: [https://academicworks.cuny.edu/gc\\_etds/1410](https://academicworks.cuny.edu/gc_etds/1410)

Discover additional works at: <https://academicworks.cuny.edu>

---

This work is made publicly available by the City University of New York (CUNY).

Contact: [AcademicWorks@cuny.edu](mailto:AcademicWorks@cuny.edu)

THE QUALITATIVE CHARACTER OF SPATIAL PERCEPTION

by

DOUGLAS B. MEEHAN

A dissertation submitted to the Doctoral Faculty in Philosophy in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

2007

UMI Number: 3249911

Copyright 2007 by  
Meehan, Douglas B.

All rights reserved.

UMI<sup>®</sup>

---

UMI Microform 3249911

Copyright 2007 by ProQuest Information and Learning Company.  
All rights reserved. This microform edition is protected against  
unauthorized copying under Title 17, United States Code.

---

ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

© 2007

DOUGLAS B. MEEHAN

All Rights Reserved

This manuscript has been read and accepted for the  
Graduate Faculty in Philosophy in satisfaction of the  
dissertation requirement for the degree of Doctor of Philosophy.

Galen Strawson

---

---

Date

---

Chair of Examining Committee

Peter Simpson

---

---

Date

---

Executive Officer

Jonathan Adler

---

Martin Davies

---

David M. Rosenthal

---

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

## Abstract

## THE QUALITATIVE CHARACTER OF SPATIAL PERCEPTION

by

Douglas B. Meehan

Advisor: Professor David M. Rosenthal

Ordinary perceiving relies heavily on our sensing the spatial properties of objects, e.g., their shapes, sizes, and locations. Such spatial perception is central in everyday life. We safely cross a street by seeing and hearing the locations of oncoming vehicles. And we often identify objects by seeing and feeling their distinctive shapes.

To understand how we perceive spatial properties, we must explain the nature of the mental states figuring in spatial perception. The experience one has when seeing a cube, e.g., differs from the experiences one has when seeing other shapes, e.g., spheres and pyramids. We must explain how such experiences differ to fully understand how we perceive differences in the spatial properties of objects. This presents a challenge often overlooked in philosophy and cognitive science. Whereas we can differentiate physical objects by their spatial properties, we cannot differentiate the experiences involved in perception in respect of their own spatial properties. Experiences are mental states, not physical objects, so they do not themselves have spatial properties; a visual

experience of a 50 ft. tall cube, e.g., isn't itself 50 ft. tall or cubical. So we must differentiate our perceptual experiences of those objects some other way, in terms of their own properties.

I argue the experiences figuring in spatial perception have mental properties distinct from, but analogous to, the spatial properties we perceive. The experience one has when seeing a square, e.g., has a property that resembles and differs from other such mental properties in ways parallel to the ways physical squares resemble and differ from other shapes. Just as squares are more similar to rectangles than triangles, the mental property of an experience of a square is more similar to that of an experience of a rectangle than that of an experience of a triangle.

I show how this theory helps solve several problems in philosophy and cognitive science; explaining change blindness, accounting for our ability to perceive combinations of distinct properties, e.g., color and shape, and determining whether the properties of experiences pertaining to the same spatial properties in different sensory modalities are themselves the same.

## Preface

Ordinary perceiving relies heavily on our sensing the spatial properties of objects, e.g., their shapes, sizes, and locations. Such spatial perception is central in everyday life. We often identify ordinary objects, such as doorknobs, stop signs, and coffee mugs, by seeing or feeling their distinctive shapes. And we safely cross a street by seeing and hearing the locations of oncoming vehicles. Further, without the ability to perceive spatial properties it would be virtually impossible to read, or to create or appreciate artworks, such as paintings, sculptures, and films.

To understand how we perceive the spatial properties of objects, we must explain the nature of the mental states figuring in spatial perception. For example, the perceptual experience one has when seeing a cube is qualitatively different from the perceptual experiences one has when seeing other shapes, such as spheres and pyramids. So we must explain how such experiences differ from each other to fully understand how we perceive such differences in the spatial properties of objects. But explaining how we individuate such experiences presents a challenge that is often overlooked in philosophy and cognitive science. Whereas we can differentiate the physical objects we perceive by their spatial properties, we cannot differentiate the experiences involved in such perception in respect of their own spatial properties. Rather, experiences are mental states, not physical objects, so they do not themselves

have spatial properties; a visual experience of a 50 ft. tall cube, e.g., is not itself 50 ft. tall or cubical. So, whereas we often differentiate physical objects in terms of their spatial properties, we must differentiate our perceptual experiences of those objects some other way, in terms of their own properties.

In this dissertation, I will focus on this and several other issues surrounding the qualitative mental states involved in our perceiving the spatial properties of objects, developing a theory of the qualitative character of those mental states and showing how that theory helps solve a number of problems in philosophy and cognitive science.

In chapter 1, I argue that the experiences figuring in spatial perception have mental properties distinct from, but nonetheless analogous in a precise way to, the spatial properties we perceive, and we individuate such experiences by these mental properties. According to this view, David Rosenthal's homomorphism theory of sensing, the experience one has when seeing a square, e.g., has a property that resembles and differs from other such mental properties in ways parallel to the ways physical squares resemble and differ from other physical shapes. Just as squares are more similar to rectangles than triangles, the mental property of an experience of a square is more similar to that of an experience of a rectangle than that of an experience of a triangle.

In the remaining four chapters, I show how homomorphism theory solves a number of problems in philosophy and cognitive science, addressing literature in philosophy, cognitive and developmental psychology, and cognitive

neuroscience. In chapters 2 and 3, I show how homomorphism theory helps explain the surprising phenomenon of change blindness, whereby one fails to notice otherwise obvious changes in a visual scene, e.g., when a central figure in the scene gradually changes from green to red. In chapter 4, I argue that this theory best explains how we perceive combinations of distinct properties, e.g., when seeing objects of the same shape but different colors, or when seeing objects of the same color but different shapes. And in chapter 5, I examine whether the properties of experiences pertaining to the same spatial properties in different sensory modalities are themselves the same. For example, do visual and tactile experiences pertaining to the same shape have some common property in virtue of which they are experiences of that same shape?

There are of course topics that I would have liked to address here, but haven't. I focus primarily on the nature of the qualitative states involved in our seeing and feeling the spatial properties of objects. But we perceive spatial properties in other sensory modalities as well. For instance, one hears where objects are located by hearing where the sounds they produce are coming from. And one can smell the locations of objects, e.g., when sniffing out the source of an odor in one's refrigerator. We are proprioceptively aware of the relative positions of our own limbs, and kinesthetically aware of their movements. And we feel bodily stimulation in various locations within our bodies; e.g., one can feel a sharp, stabbing pain running through one's leg, a tickle on one's right ankle, and a dull ache in one's stomach. A complete theory of perceptual experience

will account for all of these cases, and a theory of the qualitative character of spatial perception will help considerably in doing so. Further, a theory of perception must explain how one perceives combinations of distinct properties that one cannot perceive in the same sensory modalities. One cannot, e.g., see sound or hear color, but one can perceive a bird as being both green and melodious. Arguably, it is because one can both see and hear the location of the color and sound that one can perceive such intersensory combinations of properties. So a theory of the mental states involved in spatial perception presumably can help explain such intersensory integration. But these topics will have to wait to be addressed another time.

## Acknowledgements

I am grateful to a number of people for their contributions to this work.

David Rosenthal was an incredibly generous, demanding, and encouraging dissertation advisor. I simply would not have produced this work without his tremendous support, and his relentless demand for rigor. Most of what I have learned in graduate school I have learned from David, and I will always be indebted and grateful to him. I would also like to thank David for establishing, organizing, and running the CUNY Cognitive Science Symposium, which was central to my graduate education.

I am also incredibly grateful to the other members of my dissertation committee. Martin Davies provided detailed suggestions and challenges, and invaluable support, guidance, and advice, most from far away Canberra. Jonathan Adler, Barbara Montero, and Galen Strawson provided excellent and stimulating challenges both before and during my defense. I am also grateful to John Greenwood, an original member of my committee who was on sabbatical during the semester of my defense, and who read my materials and provided a great amount of support throughout.

I owe a great debt of gratitude to several fellow graduate students who have made my graduate experience enjoyable, and who have had a crucial impact on my work and education. I'd especially like to thank Roblin Meeks,

Josh Weisberg, Bill Seeley, Jared Blank, Peter Langland-Hassan, David Pereplyotchik, Russell Marcus, Carrie Figdor, Mark McEvoy, and Richard Brown.

Several others have also provided valuable input, through conversations and comments, and through their own works: in particular, Valtteri Arstila, Austen Clark, Diego Fernandez-Duque, Brian Glenney, Uriah Kriegel, Pete Mandik, Zenon Pylyshyn, and Peter Ross.

This work was written with the generous funding of the Mario Capelloni Foundation and The CUNY Graduate Center. I am also grateful to the National Science Foundation, the Mind-Science Foundation, and the CUNY Graduate Center for travel grants that enabled me to present much of this work at several conferences.

I have presented parts of this dissertation at several conferences, including meetings of the Association for a Scientific Study of Consciousness, the McDonnell Project in Philosophy and the Neurosciences, the Society for Philosophy and Psychology, Towards a Science of Consciousness, the Cognitive Science Society, the Brown University Graduate Student Philosophy Conference, the University of Western Ontario Graduate Student Philosophy Conference, and the New Jersey Regional Philosophy Association. I'm grateful to the organizers, commentators, and audiences of all of these events.

Finally, I am thankful to my family. My mother has always been a source of undying support and encouragement, and she has shown me that hard work, dedication, and determination do pay off. My brother, Paul, has always been

encouraging and understanding throughout this long process. I thank my mother-in-law and father-in-law, Young and Nam Suh, for their great support, encouragement, advice, and patience. And above of all, I would like to thank my wife, Caroline, who expects and deserves the best from me. Through her immeasurable patience, understanding, support, sacrifice, and love, she has made as great a contribution to this work as I have.

For Caroline

## Table of Contents

List of Figures	xvi
Chapter 1: The Qualitative Character of Spatial Perception	p. 1
§1: Introduction	p. 1
§2: Mental Qualities and Mental Space	p. 7
§3: Dennett's Concern about Mental Spatial Qualities	p. 9
§4: Homomorphism Theory	p. 17
§5: Clark's Feature-Placing Alternative	p. 38
§6: Why We Need Mental Spatial Properties, and Why Clark Does Too	p. 41
§7: Peacocke's Argument for Nonrepresentational Mental Qualities	p. 47
§8: Nonconceptual Representational Content	p. 60
§9: Homomorphism Theory and Sensory Representation	p. 67
Chapter 2: Change Blindness, Part 1	p. 80
§1: Introduction	p. 80
§2: Experiments on Change Blindness	p. 82
§3: Change Blindness and Visual Representations	p. 98
§4: Sparse Visual Representations	p. 106
§5: Dissociations of Visual Perception and Action	p. 112
§6: The Perspectival Character of Seeing	p. 122
§7: Homomorphism Theory and Sparse Sensations	p. 136
Chapter 3: Change Blindness, Part 2	p. 139
§1: Introduction	p. 139
§2: Change Blindness Despite Detailed Visual Sensations	p. 140
§3: Verbal Reports and Change Blindness: Dretske	p. 142
§4: Unconscious Change Perception During Change Blindness	p. 148
§5: Experiments on Unconscious Change Detection During Change Blindness	p. 153
§6: Neural Evidence for Change Perception During Change Blindness?	p. 179
§7: Homomorphism Theory and Change Blindness	p. 187

Chapter 4: Feature Binding and Multiple-Object Tracking	p. 195
§1: Introduction	p. 195
§2: Homomorphism Theory and the Many-Properties Problem	p. 197
§3: Experimental Support for Clark's Location-Based View	p. 201
§4: Pylyshyn's Object-Based View	p. 208
§5: Pylyshyn's Object-Based View: Object-Based Attention	p. 209
§6: Objections to Pylyshyn: Binding Without Attention	p. 214
§7: Multiple-Object Tracking	p. 225
§8: What About Represented Proximity?	p. 228
§9: Visual Indexes Are Unmotivated and Problematic	p. 231
§10: Visual Indexes	p. 238
§11: The Problem of Tracking Despite Causal Interruptions	p. 241
§12: The Problem of Detecting Features	p. 249
§13: Vision Encodes Properties of Objects It Isn't Tracking	p. 256
§14: Problems With Clark's Location-Based Binding	p. 262
Chapter 5: The Qualitative Character of Spatial Perception Across Modalities	p. 271
§1: Introduction	p. 271
§2: Feature Conjunctions and Modality Specificity	p. 273
§3: Campbell's Argument for Amodality	p. 275
§4: An Objection to Amodality	p. 288
§5: Homomorphism Theory and Modality Specificity	p. 288
§6: Crossmodal Transfer of Shape Information	p. 291
§7: Crossmodal Transfer in Infants: Facial Imitation	p. 295
§8: Tactile-to-Visual Shape Transfer	p. 301
§9: Crossmodal Transfer in Infants: Visuo-Tactile Shape Transfer	p. 311
§10: Crossmodal Shape Recognition and Modality Specificity	p. 315
§11: Neural Tactile and Visual Representations of Shape	p. 320
Bibliography	p. 325

## List of Figures

## Chapter 2: Change Blindness, Part 1

Figure 1: Adapted from Grimes (1996)	p. 85
Figure 2: Flicker Paradigm	p. 87
Figure 3: Mudsplash Paradigm	p. 89
Figure 4: Adapted from Turatto et al. (2002)	p. 93
Figure 5: The Ebbinghaus Illusion	p. 118

## Chapter 3: Change Blindness, Part 2

Figure 1: Adapted from Russell & Driver (2005)	p. 155
Figure 2: Adapted from Russell & Driver (2005)	p. 157
Figure 3: Adapted from Fernandez-Duque & Thornton (2000)	p. 163
Figure 4: Adapted from Fernandez-Duque & Thornton (2000)	p. 165
Figure 5: Adapted from Thornton & Fernandez-Duque (2000)	p. 172

## Chapter 4: Feature Binding and Multiple-Object Tracking

Figure 1: Illusory Conjunction Paradigm	p. 203
Figure 2: Adapted from Egly et al. (1994)	p. 211
Figure 3: Adapted from Baylis and Driver (1993)	p. 212
Figure 4: Adapted from Russell and Driver (2005)	p. 217
Figure 5: Adapted from Houck & Hoffman (1986)	p. 220
Figure 6: Adapted from Pylyshyn & Storm (1988)	p. 226
Figure 7: Untitled object-tracking figure	p. 244
Figure 8: Untitled object-tracking figure	p. 245
Figure 9: Untitled object-tracking figure	p. 246

## Chapter 1: The Qualitative Character of Spatial Perception

### 1. Introduction

Spatial experience pervades our perceptual awareness of our surroundings. We see objects of various sizes and shapes and at various locations, and we feel the shapes, sizes, and locations of objects we touch. We also feel bodily stimulation in various places in our bodies, e.g., pains in our backs and itches on our feet, and we sense the movements and positions of our own limbs. Our ability to perceive and sense the spatial properties of stimuli is crucial to countless daily activities. We navigate our environment by perceiving the locations and shapes of obstacles and landmarks, and we tend to our own bodies by feeling where they are damaged and stimulated.

Spatial experience is also integral to our perceiving objects. To see something as an individual object, one must distinguish it from other objects one sees. Individuating objects in this way depends on our experiencing them as spatially extended, bounded entities that bear spatial relations to other such entities. For example, one sees one's computer as distinct from the desk on which it rests. Distinguishing the computer from the desk in this way depends on one's seeing the different shapes of those objects and seeing them as occupying distinct regions of space.

Also, since we can experience the spatial properties of objects in different sensory modalities, experiencing the spatial properties of objects arguably

enables us to perceive those objects as having properties that we can sense only in different, dedicated sensory modalities, e.g., when perceiving a cup of coffee as both brown and hot, or a bird as both green and melodious. Presumably, one perceives the cup as being both brown and hot because one sees a brown cup-shaped object at the same location where one feels a hot cup-shaped object. And presumably one perceives the bird as both green and melodious because one hears a melody coming from the same place where one sees a green bird-shaped object.

And experiencing the spatial properties of objects arguably enables us to perceive objects as existing independently of our perceiving them. We assume that objects continue to exist when we do not perceive them in part because we assume that those objects can exist at locations that are beyond the limits of our sensory modalities. And that assumption arguably depends in part on our perceiving objects as bearing spatial relations to each other that are independent of the spatial relations they bear to us.

To explain how we perceive the spatial properties of objects, we must explain the nature of the mental states involved in such perception. Common sense distinguishes between two types of mental states involved in our perceptual and sensory experiences. On the one hand, perceiving involves intentional states, such as perceptual beliefs, about the objects we see, feel, and hear. Normally, when one sees an apple, e.g., one believes there is an apple present. When one feels heat emanating from a stove, one believes the stove is

hot. And when one feels a pain in one's foot, one believes that one's foot has been damaged in some way. But common sense also holds that perceiving involves qualitative states that are distinct from the intentional states involved in perceiving. In this dissertation, I address a number of problems in philosophy and cognitive science surrounding the nature of the qualitative states involved in our perceiving the spatial properties of objects.

The philosophical literature on perceptual experience is rife with debate about the qualitative character of perceiving. But this debate tends to focus primarily on the qualitative character of perceiving such properties as the colors of objects, largely ignoring the qualitative character of perceiving the spatial properties of objects. The perceptual experiences one has when seeing colors have a certain phenomenological or qualitative character that makes them seem radically different from the neurological states science supposes them to be. Philosophers also argue that bodily sensations such as pains resist scientific explanation, since they too seem so different from neurological states. Given that sensations of colors and pains seem so different from states of the brain, it is difficult to see how experiences of color and pain could result from activity in one's brain. To give a complete account of the mind, we must explain why this is so, and what this difficulty shows about the nature of such perceptual experiences. Problems related to this issue are widely referred to as the hard problem of consciousness (Chalmers, 1996; Strawson, 1994) and the problem of the explanatory gap (Kripke, 1980; Levine, 1983, 2001).

But explaining the qualitative character of seeing colors and feeling pains will not constitute a complete philosophical account of the qualitative character of perceptual and sensory experience. The experiential states involved in our perceiving the spatial properties of objects, properties such as their shapes, sizes, orientations, locations, and movements, also have distinctive qualitative characters. Just as seeing green is qualitatively different from seeing red, seeing a square is qualitatively different from seeing a triangle. And just as feeling a sharp pain is qualitatively different from feeling a dull, throbbing pain, feeling a pain in one's left knee is qualitatively different from feeling a pain in one's right shoulder. So to fully explain the qualitative character of perceptual and sensory experience, we must explain the qualitative character of perceiving the spatial properties of stimuli. We must explain the qualitative character of the mental states involved in perceiving the shapes, sizes, orientations, and locations of objects and other stimuli, such as bodily conditions.

To explain the qualitative character of perceiving spatial properties, we must explain the nature of the properties of the qualitative states, or sensations, involved in perceiving. The visual sensation one has when one sees a square is qualitatively different from the visual sensation one has when one sees a triangle. So the visual sensation one has when one sees a square has some property that the visual sensation one has when one sees a triangle does not have, and conversely the visual sensation one has when one sees a triangle has some property that the visual sensation one has when one sees a square does not

have. Likewise, the visual sensation one has when one sees a 40 ft. tall tree is different from the visual sensation one has when one sees an 80 ft. tall tree. So these visual sensations differ in respect of some properties that pertain to the different sizes of the trees one sees. And the visual sensation one has when one sees something off to one's right is qualitatively different from the visual sensation one has when one sees something off to the left. So these sensations differ in respect of some properties that pertain in some particular way to the different locations of the objects one sees.

But visual sensations are mental states, not physical objects. So they presumably do not have the same spatial properties as the objects they enable us to see. For example, a visual sensation of a 40 ft. tall tree is not itself 40 ft. tall or shaped like a tree. So we must explain the nature of the properties in virtue of which we individuate such sensations. And we must explain how those properties relate to the perceptible spatial properties of objects in a way that helps explain how we perceive those perceptible spatial properties.

This problem of explaining the qualitative character of the sensations involved in our perceiving the spatial properties of objects is of course not limited to visual cases. Feeling a cube is qualitatively different from feeling a sphere, and feeling a small sphere is qualitatively different from feeling a larger sphere. Likewise, as I noted above, feeling a pain in one's left knee is qualitatively different from feeling a pain in one's right shoulder. And feeling a sharp, localized pain in one's left thigh is qualitatively different from feeling a pain

throughout one's entire left thigh. But, like visual sensations, tactile and bodily sensations are mental states, so they do not have the same spatial properties as the tactile stimuli and bodily conditions that cause those sensations. So we must explain the nature of the properties in respect of which we individuate such qualitative mental states, or sensations.

In this chapter, I will examine various attempts to explain the qualitative character of the states in virtue of which we sense the spatial properties of stimuli. I will begin by outlining a basic view, according to which those states have mental qualities, mental properties that determine qualitative character, that pertain in some way to the nonmental, perceptible properties of objects. I will then examine Daniel Dennett's (1981) discussion of how to determine whether perceptual states do in fact have such mental qualities. I then discuss David Rosenthal's (1991, 1999, 2001, 2005) homomorphism theory of qualitative character, according to which perceptual states have mental qualities that represent perceptible spatial properties by way of homomorphisms between families of mental qualities and families of spatial properties. I then defend homomorphism theory against Austen Clark's (2000) claim that we need not countenance mental qualities that represent the spatial properties of objects to explain the qualitative character of perceiving those spatial properties. Finally, I discuss Christopher Peacocke's (1983) argument that we must countenance nonrepresentational properties of perceptual experiences in order to account for the qualitative character of perceiving spatial properties. I'll argue that we need

not commit to such nonrepresentational properties, and that homomorphism theory offers the best explanation of the qualitative character of spatial perception.

Throughout this dissertation, I further develop homomorphism theory by showing how it helps solve a number of important problems in philosophy and cognitive science surrounding the qualitative character of perceiving the spatial properties of objects. I focus primarily on vision, since it is the sensory modality most widely studied in philosophy, psychology, and neuroscience. But in chapter 5, I discuss the relations between the properties of sensations pertaining to the spatial properties of objects in different sensory modalities, focusing on the relations between seeing and feeling the same shapes. However, the theory I develop throughout the dissertation arguably applies to other sensory modalities, including those in virtue of which we are aware of the spatial properties of our own bodies, though I do not discuss such cases at length here.

## 2. Mental Qualities and Mental Space

Since the qualitative sensory states, or sensations, involved in perception are mental states, they are to be individuated in respect of their distinctly mental properties. So the sensations involved in our perceiving the spatial properties of objects are arguably to be individuated in respect of mental properties pertaining in some way to those spatial properties. On a basic view of sensing, a visual sensation of a square, e.g., has some mental property, or mental quality, in virtue

of which that sensation is a sensation of a square and not of some other perceptible shape. And a visual sensation of a triangle has some other mental quality in virtue of which it is a sensation of a triangle and not a sensation of some other perceptible shape. Likewise, a visual sensation of something off to the left has some mental quality pertaining to a region of space off to the left within one's field of view, the space in front of one's open, functioning eyes in which physical, visible objects are located. And a visual sensation of something off to the right has some mental quality pertaining to a location off to the right in one's field of view. Mental qualities are those properties of sensations that determine their qualitative characters; it is in virtue of a sensation's having the mental qualities it has that that sensation is a sensation of some particular perceptible properties and not of others.

On this basic view, for any spatial property one senses, one's sensation has a mental quality pertaining in some way to that spatial property. Such mental qualities account for the introspectible qualitative character of perceiving the spatial properties of objects. But they also account for our ability to sense those spatial properties. If our sensations did not have properties pertaining in some way to the perceptible spatial properties of objects, those sensations could not play a role in our perceiving the spatial properties of objects.

If one's sensations have such mental qualities pertaining to the perceptible spatial properties of objects, properties such as their shapes, sizes, locations, orientations, and movements, then there is arguably a mental analogue of

perceptible space. Just as there is a region of space in which perceptual stimuli are located, there is a mental analogue of that physical, perceptible region of space, and one's sensations of stimuli are in some way located there. If so, we must explain the nature of this mental space and the nature of the constituent mental properties and relations that pertain to the perceptible spatial properties of objects and the perceptible spatial relations those objects bear to each other.

### 3. Dennett's Concern about Mental Spatial Qualities

But Dennett (1978) challenges the view that visual perception involves mental states, such as sensations, with mental qualities pertaining to the spatial properties of objects. Dennett argues that the commitment to such states and properties rests on a mistaken view about what our first-person, introspective access to the states involved in perception reveals about those states.

According to Dennett, there are two distinct ways to go about determining the nature of the mental states involved in visual perception. One way, which Dennett calls the phenomenological approach, relies on one's first-person, introspective access to one's mental states. It is widely held that when one introspects the mental states involved in visual perception, one is aware of oneself as having states that are best described as mental images of what one sees.<sup>1</sup> If there are mental images, they are of course not literally images.

---

<sup>1</sup> The debate over the existence of mental imagery extends beyond a debate about the nature of the qualitative states involved in ordinary perception. Stephen Kosslyn (1994) has argued that we can best explain certain spatial

Images represent objects in virtue of resembling those objects, e.g., in respect of their spatial properties. But mental states do not themselves resemble perceptible objects; they do not have the same properties as those objects. So, presumably, committing to introspectible mental images commits us to the existence of introspectible mental states that represent objects in some way that is analogous to the way images represent objects. And, arguably, to be image-like in any relevant way, a mental state must have properties analogous to the spatial properties of objects. If so, the claim that introspection makes one aware of mental images involved in visual perception is at a minimum the claim that introspection presents us with mental states that have mental qualities pertaining in some way to the spatial properties of visual stimuli. According to the phenomenological approach to mental images, introspection reveals that we have such mental states.

The other approach to determining the nature of the mental states involved in visual perception does not commit to the existence of image-like mental states, or mental states with mental qualities analogous to the spatial properties of objects. Rather, according to this approach, which Dennett calls the scientific approach, the mental states involved in perception are theoretical

---

reasoning tasks, such as the so-called mental-rotation task, in terms of our having image-like mental states. And Zenon Pylyshyn (2003) has argued that we need not commit to such image-like states to explain our performance on these tasks. This debate is interesting, and warrants further examination. However, it is not clear how the states involved in performing such tasks as mental rotation relate to the mental states involved in ordinary perception. And I will not address these issues in this dissertation.

posits, posited to explain the typical effects of perception. Though one might claim that we are conscious of these mental states as being image-like, or as having properties analogous to the spatial properties of objects, the scientific approach individuates such states only in respect of their typical causes and effects. According to the scientific approach, such mental states are those states that are normally caused by such-and-such sensory inputs and that normally cause such-and-such behavioral outputs and other mental states. On this functionalist view, a visual sensation of a square, e.g., is a mental state that is caused by the presence of squares in one's field of view, causes other mental states, such as one's belief that there is a square present, and causes overt behavior, including, e.g., one's reaching for the object if one has the desire to grab a square object, and one's perceptually discriminating between that object and objects of other shapes.

Another typical effect of such mental states, Dennett claims, is the belief that one is having such a mental state. And it could be that that belief represents that mental state as being like an image, or as having properties analogous to spatial properties. But the scientific approach, according to Dennett, is not committed to the truth of such beliefs about the nature of those mental states (1978, p. 187). Rather, the scientific approach leaves it open that the mental state that causes one's belief that one has an image-like mental state is not in fact image-like, i.e., that it does not have properties analogous to the spatial properties of objects or images.

Dennett argues that to determine the nature of the mental states involved in visual perception we must employ both the phenomenological approach and the scientific approach. The phenomenological approach enables us to determine what exactly one is claiming about the mental states in question when one claims to have an image-like mental state. We can thus use the phenomenological approach to collect data on how we describe the mental states involved in visual perception. However, Dennett claims, we cannot assume that those descriptions are true. To determine whether those descriptions are true, he argues, we must identify the mental states that cause the beliefs we express when uttering those descriptions. To identify those mental states, we use the scientific approach. If we discover that the descriptions we gather using the phenomenological approach are true of the states we identify by the scientific approach as the causes of the beliefs expressed by our descriptions of those states, then the mental states involved in perception are in fact image-like. However, if the states we identify using the scientific approach do not have the properties introspection seems to reveal, then those states are not image-like, i.e., they do not have properties analogous to the spatial properties of objects (1978, p. 186).

So Dennett claims that determining the nature of the mental states involved in visual perception is an empirical endeavor that treats our introspective reports of mental states as empirical data we must explain. Treating introspection as unquestionably authoritative about the nature of the mental

states involved in visual perception, Dennett warns, commits one to a problematic phenomenal space in which image-like mental states are located. According to Dennett, such a phenomenal space would be "... more transparent to cognition than ordinary physical space, yet more actual and concrete than the mere logical space in which logical constructs, possible worlds, and the like reside" (1978, p. 186). Dennett further argues that we need not commit to such a phenomenal space; "... if mental images turn out to be real, they can reside quite comfortably in the physical space in our brains, and if they turn out not to be real, they can reside, with Santa Claus, in the logical space of fiction" (1978, p. 186; emphasis in the original).

Dennett claims to remain neutral about whether the mental states involved in visual perception are in fact image-like, i.e., whether they do in fact have properties analogous to the spatial properties of objects (1978, p. 188). However, I'll argue, we can best explain the nature of those mental states in terms of the view that they do in fact have properties analogous to the spatial properties of objects. Further, this explanation depends, not only on our first-person, introspective access to those mental states, but also on the observable behavioral effects of such states. This explanation thus follows Dennett's suggestion that we employ the scientific approach to determine the nature of those mental states and their properties.

I have characterized Dennett as claiming that we can employ the scientific approach to determine whether visual perception involves mental states with

properties analogous to the spatial properties of scenes and objects, i.e., whether it involves image-like mental states. However, Dennett is perhaps best read as discussing how we can determine the format of subpersonal states involved in perception, not personal-level mental states.<sup>2</sup> On this reading, Dennett holds that our introspective reports and beliefs about the nature of our mental states are never wrong; if introspection leads one to believe that perception involves image-like mental states, then we cannot discover via the scientific approach that such mental states are not in fact image-like. Rather, on this reading of Dennett, the scientific approach could reveal only whether perception involves subpersonal states with properties analogous to the spatial properties of scenes and objects. Accordingly, Dennett assumes the scientific approach will not help explain the qualitative character of the personal-level, mental states involved in spatial perception because the scientific approach provides only third-person access to the states involved in perception, whereas we have only first-person access to personal-level, mental states.

But it is unclear that folk psychology is in fact committed to the Cartesian view that the qualitative character of perceptual states is determined wholly by our first-person access to them. Rather, it could be that folk psychology holds that qualitative character is determined by factors independent of our first-person access to perceptual states. If so, Dennett's scientific approach could reveal the nature of personal-level mental states, not just that of subpersonal states.

---

<sup>2</sup> This reading is perhaps supported by Dennett's footnote (1978, p. 189) and his treatment of these issues in Dennett (1991).

Folk psychology holds that qualitative states, i.e., sensations, play a role in perception; it is in virtue of having qualitative states that one can sense the properties of stimuli. For example, it is in virtue of having a visual sensation of a red square that one sees a red square. Visual sensations of red are normally caused by red objects in one's field of view, and they normally cause other mental states, such as the perceptual belief that a red object is present, and certain discriminatory behavior. Such perceptual roles enable us to determine what qualitative state a person has by observing that person's behavior. It could be that folk psychology is committed to the view that qualitative states are determined, not simply by first-person access to them, but by their perceptual roles. If folk psychology holds that qualitative states are in fact determined by their perceptual roles, then Dennett's scientific approach could reveal the nature of those personal-level qualitative states, not just subpersonal states involved in perception. And it is unclear why Dennett would deny that folk psychology is committed to this view of qualitative character, apart from his simply assuming that folk psychology is committed to the view that qualitative character is determined solely by first-person access to qualitative states.

Further, cases of subliminal or unconscious perception arguably support the view that qualitative states are in fact determined by their perceptual roles. Such cases suggest that one can perceive a stimulus without being aware that one is perceiving it. If one unconsciously sees a square, e.g., one will form a perceptual belief to the effect that a square is present, and one could act in ways

that reflect one's seeing the square, but there will be nothing it is like for one to see that square; what it is like for one in that situation is the same as what it is like for one when one sees no square at all. Since folk psychology is committed to the view that qualitative states play a role in perception, and since some states play those perceptual roles in cases of unconscious perception, folk psychology arguably holds that qualitative states can occur without one's being conscious of them. If so, folk psychology is not committed to the Cartesian view that qualitative states are determined by one's first-person awareness of them.

One might argue that cases of so-called subliminal or unconscious perception are degenerate cases, not cases of bona fide perception. Those cases, one might argue, do not involve personal-level qualitative states.

But the states involved in these subliminal or unconscious cases function in much the same way as conscious qualitative states, except that one is not conscious of them. So, if we can account for the qualitative character of a perceptual state in terms of its perceptual role, it isn't clear why one would deny that the states involved in such cases are personal-level qualitative states, aside from one's simply assuming that personal-level qualitative states are determined by one's first-person access to them.

I argue below that we can in fact account for the qualitative character of perceptual states in a way that does not commit to the Cartesian view that qualitative character is determined solely by our first-person access to perceptual states. I further argue that Dennett's scientific approach does in fact reveal that

personal-level, qualitative mental states involved in perception have properties analogous to the spatial properties of objects.

#### 4. Homomorphism Theory

The qualitative mental states, or sensations, involved in perception are not simply the effects of stimuli on one's perceptual systems that cause introspective beliefs that one is in such qualitative states. Rather, sensations have perceptual roles. A sensation of a red square off to the left, e.g., enables one to see a red square object off to the left. And a tactile sensation of a ball in one's hand enables one to feel that ball in one's hand. These perceptual roles of sensations provide a way to determine the nature of sensations and their properties independent of our first-person, introspective access to them. And, I'll argue, we can best explain how sensations fill those perceptual roles in terms of the view that sensations have mental qualities that are analogous in a specific way to the spatial properties of objects. However, since this view holds that mental qualities are determined independently of our first-person, introspective access to them, it does not commit to the problematic mental space Dennett warns against.

Sensations enable one to perceptually discriminate among various stimuli. For example, one sees the difference between a square and a triangle in virtue of having different visual sensations. To enable such discriminations, one's sensations of a square and a triangle must differ themselves; the sensations must have properties pertaining in some way to the different shapes of the objects one

sees. Likewise, one sees two square objects as having the same shape in virtue of one's having visual sensations that both have some property that pertains to perceptible squares. Inasmuch as these properties of sensations enable one to perceptually discriminate among various stimuli on the basis of their perceptible properties, e.g., their shapes, the properties of the sensations are representational; i.e., they carry information about the stimuli sufficient for discriminating among them.

Furthermore, we discriminate shapes in terms of the ways those shapes resemble and differ from each other. For example, squares look more similar to rectangles than to triangles. And circles look more similar to ellipses than to trapezoids. We can explain how sensations enable us to perceptually discriminate among objects on the basis of their perceptible similarities and differences in terms of the homomorphism theory of qualitative character, the theory, developed by Rosenthal (1991, 1999, 2001, 2005)<sup>3</sup>, that sensations have mental properties, or mental qualities, that represent the perceptible properties of stimuli by way of homomorphisms between families of mental qualities and families of perceptible properties. I will motivate this view with respect to color vision, and I will then explain how it extends to the qualitative character of perceiving the spatial properties of objects.

---

<sup>3</sup> Homomorphism theory has roots in Wilfrid Sellars's (1963) theory of sensing. And similar views have also been held by Nelson Goodman (1977) and Sydney Shoemaker (1975).

We discriminate the perceptible colors of physical objects on the basis of the relative similarities and differences those colors bear to each other.<sup>4</sup> For example, red is more similar to orange than it is to green, and blue is more similar to green than it is to orange. Perceptible colors form a family of properties determined by such relations of similarity and difference.

The commonsense similarities and differences among colors are supported by psychophysical experiments exploiting the method of multidimensional scaling.<sup>5</sup> In psychophysical studies of color discrimination, a subject is shown two color swatches and then asked to report whether the swatches are the same color. By this method, we can determine the so-called just-noticeable differences between colors. Determining these just-noticeable differences allows us to map out a color space, with just-noticeably different colors occupying positions next to each other.

Other psychophysical methods exploiting the intransitivity of color indiscriminability provide an even more precise structure of the color space. Suppose there are three different shades of blue, shades A, B, and C. And suppose that a subject claims that A and B match. Further, suppose that subject claims that B and C also match. This does not by itself show that A, B, and C are the same color. Rather, it is sometimes the case that the subject does not report a difference between A and B, and the subject does not report a difference

---

<sup>4</sup> Perceptible colors are reflectance properties of the surfaces of physical objects.

<sup>5</sup> Austen Clark (1993) provides a detailed explanation of these methods.

between B and C, but the subject does report a difference between A and C.

This shows that A, B, and C are in fact distinct colors, even though the subject cannot always report the differences between them. In this case, color B is said to be between A and C; A is more similar to B than to C. This method is used to construct a more accurate and precise color space, or a color solid, than that constructed by testing for just-noticeable-differences between colors. And this color space reflects the commonsense relations of similarity and difference that hold between the colors; red is positioned closer in this space to orange than to green, and blue is positioned closer in this space to green than to orange.

Psychophysical experiments exploiting such methods thus confirm the relations of similarity and difference common sense takes to hold between colors.

If we see colors in respect of their relations of similarity and difference, as suggested by both commonsense categorization of colors and psychophysical experiments, we must explain how we do so. And if we make color discriminations in virtue of having visual sensations with properties pertaining to those perceptible colors of objects, we must explain how our visual sensations enable us to see the similarities and differences between colors.

According to homomorphism theory, visual sensations of color have mental qualities that correspond to the physical perceptible colors we see in virtue of resembling and differing from each other in ways parallel to the ways those perceptible colors resemble and differ from each other. For example, just as perceptible red resembles perceptible orange more than perceptible green,

the mental quality of a visual sensation of red resembles that of a visual sensation of orange more than that of a visual sensation of green. Accordingly, visual sensations of colors have mental qualities, mental colors, that represent perceptible colors in virtue of bearing resemblance relations to each other that map onto the resemblance relations perceptible colors bear to each other.

These resemblance relations holding among the mental colors map onto the resemblance relations holding among the perceptible colors in virtue of a homomorphism between the family of mental colors and the family of perceptible colors. And a particular mental color represents a particular perceptible color in virtue of its occupying the same position in the family of mental colors that that particular perceptible color occupies in the family of perceptible colors, where those positions are determined by the relative similarities between members of the respective families of properties.

I will adopt Rosenthal's notational device of suffixing a "\*" to a color predicate to indicate reference to a mental color; e.g., red\* is the mental quality of visual sensations of the color red, and green\* is the mental quality of visual sensations of the color green.

According to homomorphism theory, mental colors, or colors\*, are theoretical posits posited to explain how we discriminate colors. According to this view, we can best explain why we see red as more similar to orange than green in terms of the claim that our visual sensations of red have a mental color, red\*, that is more similar to orange\* than green\*.

One might argue that homomorphism theory provides an insufficient account of how mental qualities represent perceptible properties. A family of properties structured in terms of the relative similarities among its member properties could be homomorphic to more than one other family of properties. So it could be that the family of mental qualities that is homomorphic to the family of perceptible colors is also homomorphic to some other family of perceptible properties, e.g., the family of perceptible sounds. Therefore, one might argue, the mental qualities described above represent not only perceptible colors but also perceptible sounds.

But homomorphism theory is not committed to the view that mental qualities represent perceptible properties solely in terms of homomorphisms between families of mental qualities and families of perceptible properties. Sensations have perceptual roles specified in terms of their normal causes and effects. Sensations of colors are normally caused by the colors of visual stimuli, not by the sounds of auditory stimuli or any other properties. So however visual sensations represent colors, they represent colors, not sounds. According to this view, sensations of colors are posited to explain our ability to sense colors; they are functionally specified as those mental states that are normally caused by colored stimuli and that normally cause other mental states, e.g., perceptual beliefs about colors, and certain kinds of behavior, e.g., reports about the colors of stimuli one sees. Homomorphism theory goes beyond such functional descriptions of sensations by accounting for how we sense those colors on the

basis of their relative similarities and differences. It thus proposes a particular mode of representation in terms of a homomorphism between a family of mental qualities and the family of perceptible colors.

One might argue that we need not countenance the mental qualities homomorphism theory posits in order to explain how we perceptually discriminate among perceptible properties such as colors. It could be that we discriminate among colors, e.g., in virtue of having functionally specified intentional states, such as perceptual beliefs, that represent colors the same way other intentional states represent things. Such a view of sensing has been offered by representationalists, such as David Armstrong (1968) and George Pitcher (1970).

However, it is unclear how such representationalist views could account for the folk psychological distinction between qualitative sensory states and intentional states, such as thoughts and beliefs, that are individuated in respect of their mental attitudes and intentional contents. Homomorphism theory, on the other hand, accounts for this distinction. According to homomorphism theory, qualitative states have mental qualities that represent perceptible properties by way of homomorphisms between families of mental qualities and families of perceptible properties, but intentional states do not represent objects or properties in this way. In fact, one can have thoughts about things, such as homes, justice, and information, that, unlike colors, do not belong to well-defined property families.

Homomorphism theory describes mental qualities in terms of how they enable us to perceptually discriminate among perceptible properties, such as colors. Since such perceptual discriminations are publicly observable, homomorphism theory provides an account according to which the qualitative character of sensing is accessible from a third-person perspective.

Nevertheless, homomorphism theory does not deny that we are sometimes aware of our qualitative states from a first-person perspective, e.g., through introspection. Homomorphism theory holds that mental qualities are introspectible, though it denies that they are determined by introspection alone.

When one introspects one's visual sensations while seeing, e.g., the colors red, orange, and green, one is aware that one is having sensations of the perceptible colors red, orange, and green; i.e., one is aware that one is having the sensations of the types that enable one to see the colors red, orange, and green. Further, one is aware of the relative similarities and differences that hold between these sensations; one is aware of one's sensation of red as being more similar to one's sensation of orange than one's sensation of green. According to homomorphism theory, this is because one introspects one's sensations in respect of their mental colors, and one does that in respect of the ways those mental colors resemble and differ from each other.

This view suggests that one is introspectively aware of one's sensations only in relation to the perceptible properties those sensations enable one to perceive, i.e., only in terms of the perceptual roles one's sensations play. This in

turn suggests that when one introspects one's sensations, one applies a theory about how one perceives, e.g., colors. But, one might argue, first-person access to one's sensations is not a matter of one's applying a theory to oneself; applying theories involves inferences, whereas introspection gives one direct, noninferential access to one's mental states. So, one might argue, homomorphism theory fails to account for qualitative character in a way that is compatible with how we introspect our sensations.

But the view that introspection does in fact make us aware of our sensations in a direct, unmediated, noninferential, theory-independent way is unmotivated. It is true that when one introspects one's sensations, one is not aware of any inferences one draws from one's folk theory of perception to the conclusion that one is having a particular sensation, e.g., a sensation of red. But, as Rosenthal (1997) argues, it could be that first-person access, such as introspection, does rest on such inferences, though one is not aware of those inferences. Further, the view that introspection makes one aware of one's sensations in a theory-independent way is of a piece with the view that the qualitative character of sensations is determined wholly by one's first-person access to them. And that view, as Dennett argues, commits one to a problematic mental space. So first-person access to one's sensations is better explained as resting on inferences of which one is not aware. Homomorphism theory is thus compatible with the way we introspect our sensations.

In addition to accounting for how we sense the colors of objects, homomorphism theory also accounts for how we sense the spatial properties of objects. We see squares as more similar to rectangles than triangles, and we see circles as more similar to ellipses than trapezoids. So we see shapes as systematically resembling and differing from each other. According to homomorphism theory, one is able to see shapes as resembling and differing from each other in these ways because one's visual sensations of shapes resemble and differ from each other in ways parallel to the ways perceptible shapes resemble and differ from each other. For example, just as squares resemble rectangles more than triangles, visual sensations of squares resemble visual sensations of rectangles more than visual sensations of triangles. Visual sensations of shape thus have mental shapes, or shapes\*. And the family of shapes\* is homomorphic to the family of visible shapes.

One can of course also feel shapes, so tactile sensations of shapes also have mental shapes, or shapes\*. And, according to homomorphism theory, the mental shapes of tactile sensations resemble and differ from each other in ways parallel to the ways tangible shapes resemble and differ from each other.

Since we both see and feel shapes, we must determine whether the mental qualities of visual sensations pertaining to the shapes of objects are the same as the mental qualities of tactile sensations pertaining to those shapes. I will discuss this issue in chapter 5, where I argue that the mental qualities of visual sensations pertaining to perceptible shapes and those of tactile sensations

pertaining to the same perceptible shapes are in fact distinct, modality-specific mental qualities.

We discriminate objects on the basis of their locations as well. For example, one sees the difference between a square off to one's left and a square off to one's right. In order to enable one to see this difference in location, the sensations involved in one's seeing this difference must themselves differ in some corresponding way. But, of course, a sensation of a stimulus off to the left is not itself off to the left, and a sensation of a stimulus off to the right is not itself located off to the right. So we must explain how those sensations differ, and we must explain how the difference between such sensations relates to the difference in location between the two stimuli. This is of course an instance of the more general problem of explaining the qualitative character of spatial perception.

The problem is amplified by considering that when one has a visual sensation of a stimulus off to the right, e.g., one's sensation has mental qualities pertaining to properties other than just its location. For example, when one sees a Coke can off to the right, one sees a red cylinder off to the right. And one does so in virtue of having a sensation with mental qualities pertaining to both the color red and the cylindrical shape of the can. But the mental qualities red\* and cylindrical\* are not themselves located off to the right where the Coke can is located; they are mental properties. However, the sensation one has when seeing a red cylinder off to the right is different from the sensation one has when

seeing a red cylinder off to the left. If those sensations did not differ, they could not enable one to see the difference in location between the two stimuli. So we must explain the nature of the properties of sensations pertaining to the locations of stimuli in a way that avoids locating other mental qualities of those sensations in the field of view where the stimuli are located.<sup>6</sup>

Homomorphism theory explains the nature of the properties of sensations pertaining to the locations of stimuli without locating mental qualities, such as those pertaining to color and shape, at the same locations as those stimuli. And homomorphism theory accounts for those properties of sensations in a way that explains how they enable one to sense the locations of stimuli.

According to homomorphism theory, sensations have mental qualities that correspond to the locations of distal stimuli. A visual sensation of a square at the center of one's visual field, i.e., the region of space in front of one's open, functioning eyes in which visible objects are located, is normally caused by a square stimulus at the center of one's visual field. Stimuli in the center of the visual field normally cause sensations that have the mental quality of being at-the-center-of-the-visual-field\* (CVF\*, hereafter). And stimuli off to the right in one's visual field normally cause sensations that have the mental quality of being off-to-the-right\*.

The sum total of location\* qualities of visual sensations at a given time

---

<sup>6</sup> Frank Jackson (1977, p. 103) argues that the mental qualities of color sensations are in fact located at the same locations as the objects we see. However, we need not commit to this counterintuitive claim, as I will argue below.

constitute the mental visual field at that time. So the CVF\* is that location\* equidistant\* from each pair of opposing points on the boundary\* of the visual field, where the boundary\* is defined by the limits of locations\*. For instance, the left\* boundary is fixed by the sensation to which no other sensation is to-the-left\* of.

According to homomorphism theory, locations\* in the mental visual field correspond to locations of objects in the perceptible visual field in virtue of resembling and differing from other locations\* in ways parallel to the ways locations in the perceptible visual field resemble and differ from each other. Two stimuli can resemble each other more than either resembles a third stimulus with respect to location in the perceptible visual field. For example, two objects off to the left in one's visual field are more similar to each other than either is to an object off to the right, with respect to at least one dimension of location. Both of the objects off to the left have the property of being off to the left in one's visual field, while the third object has the property of being off to the right in one's visual field. And the two objects off to the left are more similar to a fourth object located directly in front of one than they are to the object on the right. This is because being off to the left in one's visual field is more similar to being directly in front of one than it is to being off to the right in one's visual field, at least with respect to the horizontal axis of locations within one's visual field. Likewise, visual sensations resemble and differ with respect to mental location. When one sees, e.g., a square off to the left, a triangle in the center of the visual field, and a circle

off to the right, one has a square\* sensation off-to-the-left\*, a triangular\* sensation in the CVF\*, and a circular\* sensation off-to-the-right\*. In this case, the square\* sensation resembles the triangular\* sensation more than it resembles the circular\* sensation with respect to location\*. According to homomorphism theory, this is because the mental quality to-the-left\* is more similar to the mental quality CVF\* than it is to the mental quality to-the-right\*.

The various locations\* within the mental visual field thus form a quality family of locations\* that is homomorphic to the family of locations in one's visual field. A sensation's having a particular location\* is a function of its having a mental quality the identity of which is determined by a position within this quality family of locations\*. Just as one often describes a stimulus in respect of its location relative to another stimulus, e.g., when describing a red square as being to the left of a green triangle, one can describe a sensation in respect of its location\* relative to another sensation. For example, when looking at a red square to the left of a green triangle, one's red\*, square\* sensation is to-the-left\* of one's green\*, triangular\* sensation. Such relative locations\* are not of course independent of the locations\* specified relative to the family of locations\*, i.e., those specified relative to the boundaries\* of the mental visual field. Rather, a sensation is to-the-left\* of another sensation in virtue of its having a location\* that is more similar to the location\* that defines the left\* boundary\* of the mental

visual field than it is to the location\* of the other sensation.<sup>7</sup>

Location\* properties help explain how one's having, e.g., a CVF\* sensation enables one to locate a distal stimulus directly in front of one. CVF\* sensations carry information to the effect that there is something directly in front of one's eyes, and they do so because CVF\* is the mental counterpart of being in the center of one's visual field. It is in virtue of this counterpart relation that having a CVF\* sensation helps one locate an object in the center of one's visual field, as opposed to one off to the left. And it is important to note that homomorphism theory explains the relation between locations\* and perceptible locations in terms of relations of similarity and difference that are readily accessible to us in

---

<sup>7</sup> One might worry that this account of locations\* precludes cases in which one has a single visual sensation of two stimuli, one to the left of the other. Locations\* as I have described them are properties of sensations. So, one might argue, whereas homomorphism theory can account for one's having a single sensation to-the-left\* of another sensation, it fails to account for one's having a single sensation of one stimulus to the left of another stimulus, e.g., a single sensation of a red square to the left of a green triangle. However, homomorphism theory can account for such a case in terms of one's having a single sensation composed of two parts, a red\*, square\* part that is to-the-left\* of a green\*, triangular\* part. In this case, locations\* are properties of parts of sensations.

One might also argue that homomorphism theory is committed to higher-order relations between locations\*. Just as a square can be to the left of a triangle, a location can be to the left of another location. So, one might argue, just as a sensation of a square can be to-the-left\* of a sensation of a triangle, a location\* can be to-the-left\* of another location\*.

But homomorphism theory is not in fact committed to further mental qualities of locations\*. Locations\* are determined by their relations of similarity and difference. For a particular location\* to be to-the-left\* of another location\* is for it to be more similar to the location\* that determines the left\* boundary of the mental visual field, i.e., the entire family of locations\*, than it is to the other location\*. We need not posit further, higher-order locations\* of locations\*.

ordinary visual perception and introspection.<sup>8</sup>

When one introspects one's sensations, one picks them out by their mental qualities. When one introspects one's sensation of a red square to the left of another red square, one is aware of the two sensations of red in respect of their different locations\*; one is introspectively aware that one's sensations differ in ways pertaining to the different locations of the stimuli one sees. That is, one picks out those sensations in virtue of the ways they resemble and differ from each other and other sensations.

So homomorphism theory provides an explanation of the qualitative character of the mental states involved in our perceiving the spatial properties of objects. But this theory explains such qualitative character in terms of mental qualities that are posited to explain, not just the properties of sensations as we are conscious of them, e.g., when introspecting them, but how we perceptually discriminate among stimuli on the basis of the spatial properties of those stimuli. According to homomorphism theory, perception involves mental states with mental qualities pertaining to the spatial properties of objects by way of homomorphisms between families of mental qualities and families of perceptible

---

<sup>8</sup> If visual sensations have mental qualities pertaining to the perceptible locations of visual stimuli, then we need not commit to Jackson's view that the other properties of those sensations, such as their colors\* and shapes\*, are located in the distal visual field where the stimuli they enable us to perceive are located. Rather, when one sees a Coke can, one has a sensation with mental qualities pertaining to the color, shape, size, and location of the can, i.e., it has a color\*, shape\*, size\*, and location\*. And the color\*, shape\*, and size\* of that sensation are not located in the distal visual field. Rather, they are located\* within the mental analogue of that distal visual field, the space determined by the boundaries\* of all mental qualities of visual sensations one has at that time.

spatial properties. So perception involves mental states with properties analogous to the spatial properties of the objects they enable us to perceive.

It is important to be clear that homomorphism theory is not committed to the view that one senses a stimulus in virtue of one's first sensing a sensation with mental qualities pertaining to the properties of the stimulus. For example, one does not see a red square off to the left in virtue of sensing a visual sensation with the mental qualities red\*, square\*, and off-to-the-left\* and then inferring that a red square off to the left is causing that sensation. That view, held by sense-datum theorists (e.g., Russell, 1912), requires a further explanation of how one senses one's sensations in the first place. If one must first sense one's sensations in order to see a stimulus, then presumably one has higher-order sensations with properties pertaining to the mental qualities of the visual sensation of the stimulus. But this view leads to a regress of sensations and mental qualities. Sense-datum theories avoid that regress by committing to the problematic view that one has immediate, direct acquaintance with one's sensations. That view leads to the commitment to the problematic phenomenal space Dennett warns against. And we need not commit to such acquaintance with our sensations.

According to homomorphism theory, one senses stimuli in virtue of having sensations with mental qualities that pertain to the properties of those stimuli, not in virtue of sensing those sensations. So we need not explain how one senses one's sensations to explain how one's sensations enable one to sense stimuli.

Again, according to homomorphism theory, mental qualities are theoretical posits posited to explain how one senses the perceptible properties of stimuli.

Homomorphism theory avoids countenancing the kind of phenomenal space Dennett warns against. Dennett warns that committing to a mental analogue of visible space on the basis of our first-person access to the mental states involved in perception commits one to a problematic phenomenal space that is more transparent to cognition than the physical space of objects. If we hold that the way we are conscious of our sensations determines the nature of our sensations, and if we are conscious of our sensations as having mental analogues of spatial properties, then we are committed to the view that the natures of those sensations and their mental spatial properties are given by first-person access. The physical, perceptible space in which perceptible stimuli are located is of course not given in perception; rather, things in that space are not always as they appear to be.

But, since homomorphism theory posits mental qualities to explain observable perceptual discriminations, not just to explain the way one is conscious of one's perceptual states, it is not committed to the view that the mental qualities of sensations pertaining to the spatial properties of objects are more transparent to cognition than the physical space of physical objects. So homomorphism theory is not committed to the problematic phenomenal space that Dennett warns against. According to homomorphism theory, we determine the spatial\* qualities of one's sensations by observing the perceptual

discriminations one makes. And when one is conscious of one's own sensations, e.g., when one introspects them, one is conscious of them as having mental qualities that enable one to discriminate objects on the basis of the spatial properties of those objects.

In fact, because homomorphism theory accounts for the qualitative character of perception independently of how we are conscious of the qualitative states involved in perception, it accounts for cases in which one perceives stimuli without being aware of the sensations and mental qualities in virtue of which one does so, e.g., in cases of unconscious perception.<sup>9</sup> Examples of unconscious perception are provided by psychological experiments using methods such as masked priming. In masked-priming experiments, subjects are briefly presented with a stimulus, and then a visible pattern, called a pattern mask, appears where the stimulus was located. When the stimulus and the pattern mask are presented in quick enough succession, the subject is unable to report the stimulus. However, these experiments show that such masked stimuli sometimes affect subjects' subsequent behavior, indicating that the subjects did in fact see those stimuli, even though they were unaware that they did, i.e., even though they did not consciously see them.

In one such experiment, Anthony Marcel (1983) tested for such unconscious perception using the Stroop effect. The Stroop effect is an effect on the speed at which subjects report the color of a stimulus when that stimulus is

---

<sup>9</sup> I will discuss several cases of unconscious perception in chapters 3 and 4.

accompanied by a color word referring to a different color than that of the stimulus. For example, subjects are slower at reporting the color of a red rectangle when it appears with the word 'blue' printed on it than they are when the red rectangle appears with the word 'red' printed on it. In Marcel's experiment, subjects were briefly presented with a color word, e.g., 'red', 'blue', or 'yellow', printed on a colored rectangle, e.g., a red, blue, or yellow rectangle. That stimulus was followed by a pattern mask that prevented the subject from being able to report the color word.<sup>10</sup> In some trials, the color word and the color of the rectangle were consistent; e.g., the word 'blue' was presented on a blue rectangle. In other trials, the color word and the color of the rectangle were inconsistent; e.g., the word 'blue' appeared on a red rectangle. Subjects were instructed to report the color of the rectangle as fast and as accurately as possible, and Marcel recorded their response times.

Marcel found that subjects were significantly faster at reporting the color of a rectangle when its color was consistent with the color word than when the color of the rectangle was inconsistent with the color word. So subjects' response times were sensitive to the Stroop effect, even though those subjects did not consciously see the color word, i.e., even though the subjects were unable to report the color word. That subjects' response times reflected the Stroop effect suggests that the subjects did in fact see the color words presented on the

---

<sup>10</sup> Prior to the trials, Marcel calibrated the display to ensure that the pattern mask did in fact prevent the subject from reporting the word. To do so, Marcel adjusted the interval between the stimulus presentation and the pattern mask until subjects' reports of the words were at chance levels.

colored rectangles. However, the pattern masks prevented subjects from seeing those words consciously. So we can best explain these results in terms of subjects' unconsciously seeing the color words.

Since one sees stimuli in virtue of having sensations of them, subjects in such experiments arguably have sensations of the color words, even when they cannot report those color words.<sup>11</sup> So those subjects arguably have sensations that they are not aware of having. In such cases, there is nothing it is like for one to have those sensations, but one has those sensations nonetheless.

Homomorphism theory accounts for such cases in terms of the view that sensations and their mental qualities are to be individuated in terms of their roles in perception, not in terms of the way one is conscious of those sensations and their mental qualities. So homomorphism theory explains the nature of sensations and mental qualities in a way that is compatible with cases of unconscious perception. If one can have sensations with mental qualities pertaining to the spatial properties of objects, e.g., those pertaining to the shapes, sizes, and locations of the letters of color words, without being aware that one is having such sensations with such mental qualities, then those mental

---

<sup>11</sup> One might object that such experiments do not show that subjects have visual sensations of the color words that they are unable to report. Rather, it could be that information about the color words is registered by subpersonal visual states that are sufficient for causing the Stroop Effect. But it is unclear why one would deny that subjects have visual sensations of the color words in these cases, unless one is simply assuming that if one had sensations of the color words, one would be able to report those words. Since one would arguably be able to report those words only if one consciously saw them, i.e., only if one had conscious sensations of them, this assumption begs the question against the view that sensations can occur unconsciously.

qualities are not transparent to cognition. Though homomorphism theory is committed to a mental analogue of physical, perceptible space, it is not committed to the problematic phenomenal space Dennett argues against.

Moreover, since spatial\* qualities are functional properties of mental states posited to explain perceptual discriminations, they are amenable to a physical explanation of the mind. Neural states can of course bear the relations of similarity and difference that characterize spatial\* properties. So, unlike the phenomenal space Dennett rejects, the mental spaces comprised of spatial\* properties pose no problem for a scientific theory of the mind. Sensations with mental spatial qualities, or spatial\* qualities, are in fact located in the brain; they are located wherever the neurophysiological states underlying those sensations are located.

##### 5. Clark's Feature-Placing Alternative

Austen Clark (1996, 2000) argues for a view similar to homomorphism theory to explain the qualitative character of sensing colors (1993). But he argues that sensations do not have mental qualities pertaining to the locations of stimuli, and that we do not have a mental analogue of the perceptible visual field in which visual sensations are located\*. Rather, Clark argues we need not posit mental locations of sensations to explain how we perceive the locations of objects. He offers his so-called feature-placing theory to so argue.

Clark's feature-placing theory aims to explain spatial sensing in terms of the

spatial properties of stimuli and those of sensory receptors, and in terms of patterns of neural activation that encode the spatial properties of stimuli. The only space needed to explain how we sense the locations of stimuli, according to Clark, is nonmental, physical space.

According to Clark, “[s]ensing proceeds by picking out place-times and characterizing qualities that appear at those place-times” (2000, p. 74). On his view, vision identifies a location and qualifies it as being a certain way in virtue of two distinct mechanisms. The sensation characterizes the place-time as having some property, e.g., a color, in virtue of that sensation’s having mental qualities that pertain to the properties present at that location. A sensation qualifies a place-time as being red, e.g., in virtue of having the mental quality red\*, characterized in terms of its resembling and differing from other colors\* in ways parallel to the ways the perceptible color red resembles and differs from other perceptible colors. Clark’s theory thus accords with homomorphism theory with respect to our seeing colors.

But which place-time is qualified as being red is determined, on Clark’s view, not by the location\* of one’s sensation of red, as homomorphism theory holds, but by the firing of what Clark calls a sensory name. A sensory name is a stand-in for the mechanisms of spatial discrimination. These mechanisms identify place-times by what Clark calls place-coding, which he describes with respect to somesthetic experience.

According to Clark, a group of sensory receptors on the surface of the skin

fire when they are stimulated, and they then send a neural impulse to somatosensory cortex, where a certain neural activation pattern occurs. That neural activation pattern is the neural correlate of some bodily sensation, e.g., that of an itch (2000, pp. 169-170). Where on one's body the physical itch is felt to be located, i.e., where one feels the itchy bodily stimulation, depends on which groups of sensory receptors fire (2000, p. 173). These groups of receptors are picked out by n-tuples of coordinates corresponding to the different dimensions in which the receptor groups vary in location.

Applying this view to vision, one sees a particular surface as being red, e.g., because certain receptor groups on the retina fire in a certain way, leading to a neural activation that realizes a visual sensation with the mental quality red\*. And one sees the red surface as being, e.g., off to the left because receptor groups on the right side of the retina fired, which in turn activated a sensory name, a place-coding n-tuple, encoding that receptor group. The red\* sensation realized in visual cortex is thus indexed to a place-coding n-tuple picking out that receptor group and in turn the location of the red surface.

On Clark's view, seeing something red off to the left in the visual field is a function of which receptor groups fire, and how they fire. The difference between seeing something red off to the left in the visual field and seeing something red in the center of the visual field is just a difference in which retinal receptor groups fire and what location a sensory name encodes as a result. So, Clark concludes, feature-placing explains the qualitative character of sensing without committing to

mental qualities pertaining to the spatial properties of stimuli or to a mental analogue of visible space.

#### 6. Why We Need Mental Spatial Qualities, and Why Clark Does Too

But Clark's feature-placing view fails to account for the differences between the mental states involved in our sensing stimuli at different locations. We cannot pick out our sensations of stimuli at different locations without reference to properties of those sensations that pertain to the locations of stimuli. And we cannot identify the neurophysiological processes that Clark claims enable one to sense the locations of stimuli without first picking out sensations by their mental properties that pertain to the locations of the objects we sense. Clark attempts to avoid positing mental space and mental analogues of the perceptible locations of stimuli by trying to account for sensory localization in terms of neurophysiological mechanisms and the spatial properties of stimuli and sensory receptors. But his feature-placing theory can get off the ground only if it countenances mental qualities of sensations pertaining to the locations of stimuli.

If Clark is right that sensations do not have mental qualities pertaining to the locations of stimuli we sense, then sensations of stimuli located in different places do not themselves differ in ways pertaining to the different locations of those stimuli. In this case, the sensations one has when one sees a red square above a green triangle are no different from the sensations one has when one sees a red square below a green triangle.

But inasmuch as sensations are theoretical posits posited by folk psychology to explain how we make perceptual discriminations, the sensations that one has when one discriminates between a scene consisting of a red square above a green triangle and a scene consisting of a red square below a green triangle do in fact differ in some way that pertains to the difference between the relative locations of the stimuli. Further, common sense holds that the sensations one has when seeing those two different scenes have introspectible qualitative differences. If so, the sensations presumably have mental qualities pertaining to the perceptible locations of stimuli.

According to Clark, the difference between the states involved in one's discriminating between these two different scenes is a difference in the neurophysiological sensory names that fire in conjunction with the sensations of color, it is not a difference in the mental qualities of the sensations. So Clark must accept that the sensations themselves do not differ in these two cases.

But we need not accept that conclusion. Presumably, sensations themselves are realized by neurophysiological states. The neurophysiological sensory names Clark posits could be the neural correlates of locations\*. Since common sense holds that sensations of stimuli in different locations themselves differ in ways pertaining to those differences in location, the best explanation is that sensations have mental locations, or locations\*. And, since sensory names, according to Clark, are theoretical posits posited to explain how one senses the locations of stimuli, they are posited to fill the same perceptual role that common

sense posits sensations of objects' locations to fill. Clark's feature-placing theory does not provide an alternative to the view that sensations have mental qualities pertaining to the locations of objects.

However, Clark does offer another argument against the view that sensations have mental qualities pertaining to the locations of stimuli. Adapting Frank Jackson's (1977) so-called many-properties problem, Clark argues that if one's sensations had such mental locations, one could not discriminate between two scenes consisting of different combinations of the same properties.

One can see the difference between the following two scenes:

- a) A red square at location L1 and a green triangle at location L2.
- b) A green square at location L1 and a red triangle at location L2.

Clark argues that if the sensations one has when seeing scenes (a) and (b) had mental qualities pertaining to the locations of the colored shapes, i.e., locations L1 and L2, then those sensations would be identical. The sensation one has when seeing scene (a) would have the mental qualities red\*, square\*, L1\*, green\*, triangular\*, and L2\*. And the sensation one has when seeing scene (b) would have the mental qualities green\*, square\*, L1\*, red\*, triangular\*, and L2\*. Since those sensations would have exactly the same mental qualities, they would be identical sensations. But if those sensations are identical, they could not enable one to see the difference between scene (a) and scene (b).

Clark argues that this consideration shows we do not sense the locations of stimuli in virtue of having sensations with mental locations. Rather, he claims, we sense the locations of stimuli in virtue of having sensory names that fire in conjunction with sensations of other perceptible properties, e.g., colors and shape. Sensory names, according to Clark, serve to bind distinct mental qualities in the right ways, e.g., binding red\* with square\* and green\* with triangular\* when one sees scene (a), and binding green\* with square\* and red\* with triangular\* when seeing scene (b).

But it could be that sensations have mental qualities that enable one to sense the locations of stimuli and that those mental qualities play a special role in sensing combinations of distinct properties, such as color and shape. So Clark's view that the mechanisms in virtue of which one senses the locations of stimuli play a special role in binding separate mental qualities pertaining to, e.g., color and shape is compatible with the view that those mechanisms are mental qualities.

I'll discuss this issue at length in chapter 4. There I argue that we need not commit to the view that the mechanisms in virtue of which one senses the locations of stimuli play a special role in sensing feature conjunctions. Rather, I argue, distinct mental qualities such as colors\* and shapes\* are interdependent and need not be bound by a separate mechanism.

Not only is the view that sensations have mental qualities pertaining to the perceptible locations of stimuli compatible with Clark's view that there are

neurophysiological states that enable one to sense the locations of stimuli, but we cannot identify the neurophysiological states enabling sensory localization without first identifying those states by their distinctly mental properties.

Identifying the neurophysiological state that enables one to see something in the center of one's visual field requires that we have some way of picking out that state independent of its neurophysiological properties. We do not do this just by identifying properties of sensory receptors, or by identifying neural activation patterns. Rather, we pick out the state by the role it plays in enabling one to sense stimuli, and by the properties in virtue of which that state plays that role. Since sensing is a folk psychological phenomenon, the state that plays that role is a mental state, i.e., a sensation, and we pick it out by its mental properties, specifically the mental properties in virtue of which that state enables one to sense the perceptible properties of stimuli. Since such sensations enable one to sense different locations of objects, those states have properties that pertain in some way to the locations of those objects. So picking out the neurophysiological states that enable one to sense the locations of objects is a matter of picking out a sensation in respect of its mental qualities pertaining to those locations.

We determine that firings of particular receptor groups and neural mechanisms enable one to sense the locations of objects by discovering that those firings occur whenever subjects sense the locations of objects. To determine which receptor groups are firing, we monitor neurophysiological

activity in the subjects. To determine what the subject is sensing, we monitor the subject's overt and verbal behavior. If the subject reports having a sensation of something red in the center of the visual field, e.g., we infer that the subject has a sensation of something red in the center of the visual field. If the subject exhibits some other behavior indicating that the subject sees a red stimulus in the center of the visual field, we infer that the subject has a sensation of a red stimulus there, i.e., we infer that the subject has a mental state with properties pertaining to the color and location of that stimulus. To determine what neurophysiological mechanisms underlie the subject's seeing a red stimulus at the center of the visual field, we monitor the subject's neural activity while the subject exhibits the effects of seeing that stimulus, which in turn shows that the subject is in a state in virtue of which the subject sees that stimulus, i.e., a sensation of a red stimulus in the center of the visual field. So identifying Clark's sensory names in the first place depends on our picking out the sensations in virtue of which one senses the locations of objects. Again, it is not clear why Clark thinks sensory names provide an alternative to such mental qualities.

Given that homomorphism theory accounts for our sensing the spatial properties of stimuli, and it accounts for the introspectible qualitative character of the sensations in virtue of which we do so, homomorphism theory best explains the nature of the properties in respect of which we individuate sensations of objects' spatial properties. According to this view, a visual sensation of a 40 ft. tall red square off to the left is not itself 40 ft. tall, square, or off to the left, just as

it is not red. Rather, that visual sensation has mental qualities that represent the perceptible size, color, shape, and location of that object. And those mental qualities represent those perceptible properties in virtue of resembling and differing from other mental qualities in their respective quality families in ways parallel to the ways the perceptible size, color, shape, and location resemble and differ from other perceptible properties in their respective families of properties. For example, just as perceptible squares are more similar to perceptible rectangles than perceptible triangles, square\* is more similar to rectangular\* than triangular\*.

The mental qualities homomorphism theory posits to explain the qualitative character of sensing and perception are representational properties. They are representational because they encode information about the perceptible properties of the stimuli one senses. And it is because mental qualities encode such information that they are able to fill the perceptual roles they fill.

### 7. Peacocke's Argument for Nonrepresentational Mental Qualities

Christopher Peacocke (1983) argues that an explanation of the qualitative character of perceptual experience involves more than just an account of the properties of those experiences in virtue of which one perceives stimuli. Perceptual experiences, he argues, have nonrepresentational features as well.

Peacocke cites the phenomenon of size constancy to illustrate this point.

When one sees two trees of the same height but at different distances from one, one's visual experience represents the trees as being the same height, so one perceives the two trees as being the same height. "Yet," writes Peacocke, "there is also some sense in which the nearer tree occupies more of your visual field than the more distant tree. This is as much a feature of your experience itself as is its representing the trees as being the same height" (1983, p. 12). But, Peacocke claims, "... no veridical experience can represent one tree as larger than another and also as the same size as the other" (1983, p. 12). So, he concludes, the aspect of one's visual experience whereby the closer tree occupies more of one's visual field than the farther tree is a nonrepresentational aspect of that experience.

Peacocke calls such nonrepresentational aspects of perceptual experiences sensational properties. According to this view, though one's visual experiences of the two trees represent them as being the same size, one's experience of the closer tree differs from one's experience of the farther tree in respect of some nonrepresentational sensational property. Peacocke holds that we can best explain the way in which the closer tree occupies more of the visual field than the farther tree, or looks in some way larger than the farther tree, in terms of the view that the experience of the closer tree and the experience of the farther tree differ in respect of their sensational properties.

Peacocke refers to sensational properties with primed predicates. Accordingly, the visual experience of the closer tree is large', and the visual

experience of the farther tree is small’.

This account extends to perceptual experiences of other spatial properties of objects, not just their sizes. For example, when one sees circular objects, such as pennies, tilted slightly away from one’s line of sight, one usually sees them as being circular. However, such circular objects look in some way similar to elliptical objects presented perpendicular to one’s line of sight, even though one does not see the circular objects as being elliptical. On Peacocke’s view, this is because one’s visual experience of a circular object tilted away from one’s line of sight has a sensational property, elliptical’, that one’s visual experiences of elliptical objects presented perpendicular to one’s line of sight also have.<sup>12</sup>

Peacocke’s argument that perceptual experiences have nonrepresentational properties, i.e., sensational properties, rests on the assumption that those perceptual experiences would represent contradictions if all of their properties were representational. And that assumption rests in turn on the further assumption that if one’s perceptual experiences represented the closer tree in Peacocke’s example as occupying more of one’s visual field than the farther tree, as Peacocke describes the situation, one’s experience would represent the two trees as being different sizes.

But perhaps one’s visual experience represents one tree as having the feature Peacocke describes as occupying more of one’s visual field than the other tree in virtue of representing those trees as differing in some property other

---

<sup>12</sup> Peacocke also applies this view to visual experiences of colors.

than their sizes. If so, it could be that a visual experience that represents the two trees as being the same size and as occupying different portions of the visual field does not have contradictory representational content. On such a view, we need not commit to nonrepresentational properties of visual experiences to account for such cases. But we must then determine what properties of the trees other than their sizes one's visual experiences represent such that we describe the closer tree as occupying more of one's visual field than the farther tree, or such that we claim that the closer tree looks in some way larger than the farther tree.

The two trees do in fact differ in respect of some property other than their sizes. The two trees are located at different distances from the perceiver. So, perhaps, we can best explain the way in which one tree occupies more of the visual field than the other tree, or the sense in which one tree looks in some way larger than the other tree, in terms of the visual experience's representing the trees as being located at different distances from one.

But this fails to account for the phenomenon Peacocke describes. Two trees at different distances could occupy the same amount of one's visual field, or look in some way the same size, if those trees are different sizes. And two trees located the same distance away from the perceiver could occupy different amounts of the perceiver's visual field if those trees are different sizes. So the additional aspect of visual experience that Peacocke describes in terms of one tree's occupying more of the visual field than the other tree is not simply a

function of one's seeing the two trees as being located at different distances from one.

Still, even though the additional aspects of one's visual experiences of the trees do not represent the sizes or distances of those trees, such additional aspects of visual experiences do bear nonarbitrary relations to the sizes and distances of such stimuli. One's visual experiences of two trees of the same size and at the same distance away do not differ in respect of this additional aspect. And it is by altering the sizes and distances of stimuli that we change this aspect of our experiences. So the additional aspect of visual experience that Peacocke attempts to explain in terms of sensational properties is a function of the size and distance of a stimulus, even though it is not a function of either size or distance individually.

That the additional aspects of these visual experiences bear nonarbitrary relations to the sizes and distances of stimuli is also suggested by the way we describe the additional aspects. It is arguably most natural to describe such aspects of one's experience of the two trees using size predicates, as Peacocke does. And perhaps, as Peacocke suggests, one would describe the trees themselves as in some way looking different in size, or as occupying different portions of the visual field.

However, it is important to be careful when drawing conclusions based on these ways of describing such aspects of one's perceptual experiences. The way Peacocke describes things is problematic. He claims that the closer tree

occupies more of one's visual field than the farther tree. But this is not how ordinary perceivers describe such cases. If 'visual field' is a commonsense expression, it is arguably used to refer to the three-dimensional space in which the objects one sees are located. But the two trees in Peacocke's example are the same size, so they occupy the same amounts of the three-dimensional visual field. Moreover, since one sees the two trees in Peacocke's example as being the same size, one sees them as occupying the same amount of space within such a three-dimensional visual field. And one would not claim that the trees appear to occupy different amounts of that three-dimensional space. So, 'visual field', as Peacocke uses it, is presumably a technical expression, not one that ordinary perceivers would use to describe such cases.

Perhaps ordinary perceivers would describe the closer tree as looking in some way larger than the farther tree. This way of describing the ways the trees look reveals that we are not describing them simply in terms of the sizes they look to be, but that the additional aspect we are describing is in some way related to the sizes of the trees. If in such cases we are describing nonrepresentational aspects of our experiences, as Peacocke argues, then we must explain why we use size predicates to describe them, and why we describe the trees themselves as looking different from each other. On the other hand, perhaps we describe such cases using size predicates because the additional aspects of our experiences represent perceptible properties of stimuli that relate in some nonarbitrary way to the sizes of those stimuli. If so, we must determine what

those properties are and how they relate to the sizes of stimuli.

The two trees in Peacocke's example do in fact differ in respect of properties distinct from but related to both their distances from the perceiver and their sizes. So, perhaps, one's visual experiences represent the two trees as differing in respect of those properties.

The two trees project retinal images of different sizes; the two trees subtend different retinal angles. So the trees must differ in respect of some property in virtue of which they subtend those different retinal angles. And two trees of different sizes and at different distances from one could subtend the same retinal angles, so they could project retinal images of the same size. So such trees have some property in common in virtue of which they project retinal images of the same size, even though the trees themselves are not the same size. Such properties are a function of both an object's distance from the perceiver and its size.

Perhaps one's visual experiences of the two trees represent those properties in virtue of which the trees subtend the retinal angles they subtend. Since those properties are distinct from the sizes of the trees, a visual experience could represent the trees as differing in respect of those properties without representing the trees as differing in respect of size. In that case, one's visual experience could represent the trees as differing in respect of the properties in virtue of which the trees subtend different retinal angles while also representing the trees as being the same size. If so, we need not countenance

nonrepresentational properties of visual experiences to account for the additional aspect of visual experience that Peacocke describes.

The property in virtue of which an object projects a retinal image of a certain size is a property that object has in relation to the perceiver, since the property is a function not only of the size of the object but also of its distance from the perceiver. I'll call these perceiver-relative properties P-properties, and I'll prefix a 'P' to a predicate to refer to the corresponding P-property.<sup>13</sup> For example, I'll call the property in virtue of which an object subtends a particular retinal angle the object's P-size.

On this view, when one sees two trees of the same size but at different distances from one, one might describe the closer tree as occupying more of one's visual field than the farther tree, or one might describe the closer tree as looking in some way larger than the farther tree, because one's visual experience represents the trees as having different P-sizes. And, again, the trees do in fact have different P-sizes; if they did not, they would project retinal images of the same size.

If this view is correct, then one's experience could represent the two trees as being the same size while also representing them as looking in some way different in size, or, as Peacocke describes the case, as occupying different portions of one's visual field. In this case, one's experience represents the trees

---

<sup>13</sup> Alva Noë (2004) uses the same notation to refer to such properties, which he calls perspectival properties. I'll discuss Noë's treatment of our perceiving these properties in the next chapter.

as being the same size but as having different P-sizes; one's experience represents the closer tree as P-larger than the farther tree. If so, one's visual experience does not have contradictory representational content.

But Peacocke rejects a similar account due to Irvin Rock (1975).

According to Rock, in normal perception one perceives at least two distinct but related stimuli at once. One perceives what Rock calls the distal stimulus, the object at some distance from one's eyes, and one perceives what Rock calls the proximal stimulus, the image on the retina that distal stimulus causes. On this view, when one sees the size of an object, one sees both the size of the distal stimulus and the size of the proximal stimulus, i.e., the retinal angle the distal stimulus subtends. Accordingly, one sees the two trees in Peacocke's example as being the same size because one's visual experience represents the trees as being the same size. And the closer tree in some way occupies more of the visual field than the other tree, or looks in some way larger than the other tree, not because one's visual experience represents the trees as being different sizes, but because one's visual experience represents the proximal stimuli caused by the trees as being different sizes. Since the distal and proximal stimuli are distinct, this experience would not represent a contradiction.

However, according to Peacocke, in order for one's experience to represent a property of a stimulus, one must possess the concept of that property. If Rock's view is correct, he argues, one must have the sophisticated concept of a retinal angle to have the experience one has when one sees the two trees

(1983, pp. 19-20). But, Peacocke argues, we need not have the concept of a retinal angle to have such an experience. He claims that "... for an unsophisticated perceiver who does not have the concept of subtended angle it is nevertheless true that one object takes up more of his visual field than another, just as it does for a more sophisticated theorist" (1983, p. 20). If Peacocke is right, the view that one's visual experience represents the proximal stimuli as subtending different retinal angles does not help explain the aspect whereby the two trees occupy different amounts of one's visual field, or look in some way different in size.

Nevertheless, it could be that one's visual experience represents the two trees as having different P-sizes without representing the visual angles subtended by the trees as such. The P-size of a tree is a property of the tree itself; it is thus a property of the distal stimulus, not a property of the proximal stimulus. The P-size of an object one sees is in fact the property in virtue of which that object subtends the particular retinal angle it subtends. But the P-size of an object has effects other than subtending a particular retinal angle. For example, if two objects of the same size are positioned at different distances from a mirror, the closer object will produce a larger reflection than the farther object. That is, the reflection of the closer object will occupy a larger region of the mirror's surface than the reflection of the farther tree will. Likewise, an object close to a surface will cast a smaller shadow on that surface than an object of the same size that is farther away from the surface, assuming that the light source is

located directly behind the two objects. If the two objects did not differ in respect of some property, they would not produce reflections and shadows of different sizes. It could be that when one perceives the P-size of an object, one does so in virtue of having some concept of a property with such effects, even if one does not have the concept of a subtended angle or the concept of a retina.

Further, it is widely held that for one to perceive the properties of stimuli, such as their sizes and shapes, vision must compute those properties from properties those objects have in relation to the perceiver (see, e.g., Marr, 1982).<sup>14</sup> This is most often discussed in regard to the problem of inverse optics, the problem of explaining how one sees stimuli as having invariant, perceiver-independent properties given the impoverished, perceiver-dependent nature of visual stimulation. Size and distance perception provide examples of this problem. Again, two objects of the same size but at different distances cause different visual stimulation; the closer object projects a larger retinal image than the farther object. But somehow one perceives the objects as being the same size but at different distances. In order to enable one to perceive the invariant properties of stimuli, the visual system must disambiguate the visual stimulation, e.g., to determine whether the two objects are of the same size but at different distances, different sizes and at different distances, or different sizes but at the same distance. To compute the invariant sizes of stimuli, the visual system can

---

<sup>14</sup> The view that vision computes the invariant properties of stimuli from impoverished, ambiguous stimulation has been challenged most notably by the psychologist James Gibson (1966, 1979) and more recently by Noë (2004). I'll discuss Noë's view in the next chapter.

exploit various depth cues to first determine how far away the stimuli are. Once vision computes the distance of a stimulus, it can use that information to compute its size.

One such depth cue that vision exploits to determine an object's distance is motion parallax. When one moves one's eyes while looking at an object, the retinal stimulation caused by that object also moves. And the retinal stimulation caused by an object nearby moves more with a movement of one's eyes than the retinal stimulation caused by an object further away. So the visual system can use information about changes in retinal stimulation caused by eye movements to disambiguate the retinal stimulation and determine how far away a stimulus is. Once vision determines how far away the stimulus is, it can then compute its size using principles of trigonometry. Specifically, vision could compute the height of an object as the product of the distance of that object and the tangent of the retinal angle the object subtends. Of course, one need not understand trigonometry for vision to perform such computations. Rather, these computations are performed subpersonally.

But to perform such computations, vision must form representations of the visual stimulation caused by the distal stimulus. And the properties of the retinal stimulation, as well as those of the subsequent visual representations caused by that retinal stimulation, correspond to the properties of the distal stimulus that cause them. But the properties of the visual representations do not correspond to the sizes of those stimuli and they do not correspond to the distances of those

stimuli, since stimuli of various sizes and distances can cause the same retinal stimulation. Stimuli of different sizes and distances cause the same retinal stimulation because they have certain properties in common, i.e., P-properties. So vision enables one to perceive invariant, perceiver-independent properties of distal stimuli in virtue of performing computations on representations of perceiver-dependent properties, such as P-properties.

If vision uses representations of P-properties to enable perception of invariant, perceiver-independent properties, perhaps we can best explain the aspect of visual experience Peacocke describes as an object's occupying more of one's visual field than another object of the same size in terms of one's having visual representations of the different P-sizes of those objects.

If Peacocke is right that one's perceptual experiences can represent only those properties one has concepts of, then anyone who sees the two trees in Peacocke's example as being the same size arguably has concepts of P-properties. On the other hand, perhaps the visual system represents P-properties without one's possessing concepts of those P-properties. If so, Peacocke is wrong that representing properties requires concepts of those properties.<sup>15</sup> In fact, Peacocke (1992, 2001) himself argues for such

---

<sup>15</sup> Perhaps Peacocke would claim that what I call representations of P-properties are what he would call states with informational content; i.e., states that carry information about a stimulus but that do so in a referentially transparent, nonconceptual way. Peacocke claims that he is not arguing that intrinsic, sensational properties of experience are determined by informational content (1983, p. 8). So, Peacocke might argue, our views are compatible.

nonconceptual perceptual representation.

### 8. Nonconceptual Representational Content

Common sense distinguishes between qualitative and intentional aspects of perception. Inasmuch as one can see that there is, e.g., a Coke can on the table, one's visual perception must involve an intentional state about the Coke can. And one could not have such an intentional state if one did not have a concept of a Coke can. However, common sense is arguably not committed to the view that the qualitative aspects of perceiving also require such concepts. Presumably, one can have a sensation of a Coke can without perceiving it as a Coke can. Nevertheless, common sense also holds that one's visual sensations do in fact play a role in perception. So sensations, according to common sense, do in fact represent perceptible stimuli in some way. Perhaps sensations represent stimuli nonconceptually, i.e., independent of the concepts required for one to have intentional states about those stimuli. If so, perhaps one's sensations represent P-properties independently of one's having any concepts of

---

However, Peacocke also claims that sensational properties are "... properties an experience has in virtue of some aspect—other than its representational content—of what it is like to have that experience" (1983, p. 5). Since there is something it is like to have an experience only when one is conscious of one's experience, sensational properties are determined by the way one is conscious of one's experience. But information-carrying states need not be intrinsic, conscious states. So establishing that experiences involve information-carrying states does not establish that they have intrinsically conscious properties. Even if my use of 'representational' is so broad as to include informational content, my argument still denies that Peacocke has established that experiences have sensational properties.

P-properties. In that case, it could be that the problematic difference between one's visual experience of the closer tree and one's experience of the farther tree that Peacocke discusses is a function of those experiences' nonconceptually representing the trees as having different P-sizes.

In fact, Peacocke (1992, 2001) is one of the strongest proponents of the view that perception has nonconceptual representational content in addition to conceptual content.<sup>16</sup> However, he claims that the issue of nonconceptual content is independent of the issues surrounding sensational properties, suggesting that he thinks perception has nonrepresentational sensational properties, even if it also has nonconceptual representational content.

I will first discuss several arguments for the view that perception has nonconceptual representational content. I will then argue that if perception has such nonconceptual representational content, we need not commit to the existence of nonrepresentational sensational properties. Finally, I'll argue that homomorphism theory best accounts for nonconceptual representation and for the problematic aspects of perceptual experiences that Peacocke (1983) attempts to explain in terms of nonrepresentational sensational properties.

Peacocke's arguments for the existence of nonconceptual representational content depend on his views of conceptual content. He argues that giving a noncircular explanation of what it is for one to possess a perceptual

---

<sup>16</sup> José Luis Bermúdez (1998), Tim Crane (1988a, b), Adrian Cussins (1990), Fred Dretske (1995), Gareth Evans (1982), Susan Hurley (1998), and Michael Tye (1995) also argue that perception has nonconceptual content.

concept, such as the concept of the color red or the concept of a square, requires nonconceptual content. According to Peacocke, concepts are to be individuated by their so-called possession conditions, which describe their functional roles. Perceptual concepts, such as the concept of the color red or the concept of a square, are to be individuated by their roles in the formation of perceptual judgments.

Peacocke writes:

We may individuate a perceptual concept C in part by a statement of this form: it is that concept C to possess which a thinker must be willing to judge that certain things are C in such and such circumstances in which he perceptually experiences them as falling under C... (1992, pp. 88–89)

Accordingly, the possession conditions for the perceptual concept of a square are as follows:

The perceptual concept of a square is that concept to possess which a thinker must be willing to judge that certain things are square in such and such circumstances in which he perceptually experiences them as falling under the concept of a square.

On Peacocke's view of concepts, one has the concept of a square if one can judge, or think, that something one sees is square based on one's seeing that that thing is square. But, Peacocke argues, such possession conditions for the concept of a square are circular if one must already possess the concept of a square in order to see something as square. So, he concludes, seeing something as square does not require that one possesses the concept of a square. Rather, seeing something as square involves nonconceptual representational content.

Adrian Cussins (1990) and Peacocke (2001) offer a similar argument for nonconceptual representational content. They argue that one could not acquire perceptual concepts, such as the concept of a square, if perceptual experience did not already have nonconceptual content. For example, if seeing a square does not involve a nonconceptual representation of a square, then one could not learn the concept of a square by seeing a square, since seeing a square would require that one already has the concept of a square. So if perceptual experiences, such as those involved in seeing squares, do not have nonconceptual content, then perceptual concepts, such as the concept of a square, are innate. Since such concepts are not innate, Cussins and Peacocke hold, perceptual experiences have nonconceptual content.

Peacocke also argues that if perceptual experiences did not have nonconceptual content, those experiences could not rationally justify one's

perceptual beliefs (1992, p. 80). When one sees a red square off to the left, one forms the perceptual belief that there is a red square object off to the left. And that perceptual belief is based on one's seeing such an object. If seeing that object does not involve a representation of it as a red square off to the left, then seeing the object does not rationally justify one's belief that there is a red square object off to the left. If such perceptual beliefs are in most cases rationally justified, as Peacocke assumes, perceptual experiences have nonconceptual representational content.

We can give a similar argument without committing to the rational justification of perceptual beliefs. One's perceptual beliefs are reliable, i.e., they are usually true. If one's perceptual beliefs were not usually true, one would not form them. And one's perceptual beliefs are caused by one's perceptual experiences. So one's perceptual experiences are reliable indicators of the properties of stimuli. This in turn suggests that one's perceptual experiences represent those stimuli.

Peacocke also claims that we can best explain the fine-grained nature of perceptual experience in terms of nonconceptual representational contents (1992, pp. 67-68). According to Peacocke, the amount of detail of one's perceptual experiences outstrips the conceptual content of those experiences. For example, when one describes the shapes of mountains one is looking at, one's description is much less specific than one's visual experience; one's description of the shapes of the mountains applies equally well to other mountain

ranges, but one has different visual experiences when seeing those different mountain ranges. One's perceptual experiences thus represent perceptible details that one's descriptions, and thus one's concepts, fail to represent.

Tim Crane (1988a, 1988b) offers yet another argument for nonconceptual representational content. According to Crane, certain illusions present problems for the view that the representational contents of perceptual experiences are exhausted by their conceptual contents, i.e., the contents of intentional states such as beliefs. For example, when one sees a straight stick in water, the stick looks in some way that is perhaps best described as broken or bent. However, one believes that the stick is straight and intact. So we must explain how something can look in some way broken when one believes it is intact.

Some, e.g., David Armstrong (1968) and George Pitcher (1971) claim that we can best explain such cases in terms of the view that one has distinct contradictory beliefs. When one sees the stick in water, they argue, one has both the belief that the stick is straight and intact and the belief that the stick is bent or broken.

But, Crane argues, when one discovers that one's beliefs are false, one ceases to hold those beliefs. If one believes that the stick is in fact broken, as Armstrong and Pitcher claim, one would stop believing that the stick in the water is broken when one discovers that it is intact. But discovering that the stick is not in fact broken does not eliminate the illusion; those who are aware that the stick is intact still experience the illusion. So, Crane concludes, the illusion does not

result from one's believing that the stick is broken.

Rather, Crane claims that the mental state in virtue of which the stick looks in some way broken is informationally encapsulated, in Jerry Fodor's (1983) sense; that state is unaffected by other mental states one has. And, Crane further claims, this suggests that perceptual experiences have two distinct kinds of contents, conceptual contents, which are affected by one's other mental states, and nonconceptual contents, which are not affected by one's other mental states. Accordingly, the stick looks in some way broken because one's visual experience nonconceptually represents it as broken, even though one believes that the stick is intact and one does not believe that the stick is broken.

If perceptual experiences do in fact have nonconceptual representational content, i.e., if they represent features of stimuli independent of the concepts of those features, then we can explain the problematic additional aspect of visual experience that Peacocke (1983) describes without committing to the nonrepresentational features of perceptual experiences he commits to. It could be that when one sees two trees of the same size but at different distances, one's visual experience nonconceptually represents the closer tree and the farther tree as having different P-sizes, the properties in virtue of which those trees subtend different retinal angles. In this case, we describe the closer tree as looking in some way larger than the farther tree, or as occupying more of the visual field than the other tree, because one's experiences of the trees nonconceptually represent them as having different P-sizes. If so, one need not

have the concept of a P-size in order to perceive a P-size.

One might object to the above arguments for nonconceptual representational content on a number of grounds.<sup>17</sup> For example, one might argue, against Cussins and Peacocke, that there is no reason to think that perceptual concepts are not innate. Or one might argue, against Peacocke, that one could describe a complex stimulus, such as a craggy mountain range, in such detail that one's description would in fact capture all of the visible detail of that stimulus. Nevertheless, common sense arguably holds that the qualitative states, i.e., sensations, involved in perception are not conceptual in the way that intentional states, such as perceptual beliefs, are. And common sense also arguably holds that those qualitative states play perceptual roles. So, if we can explain how sensations represent the perceptible properties of objects, we need not commit to the existence of nonrepresentational sensational properties to account for the phenomenon Peacocke describes in his example of the visual experience of the two trees.

#### 9. Homomorphism Theory and Sensory Representation

We can explain how the qualitative states involved in perception represent perceptible properties independently of one's having the concepts of those perceptible properties in terms of homomorphism theory. Again, according to homomorphism theory, sensations enable one to sense the properties of stimuli

---

<sup>17</sup> John McDowell (1994) and Bill Brewer (1999) both argue against the existence of nonconceptual perceptual content.

in virtue of having mental analogues of those properties. And those mental analogues pertain to their perceptible counterparts by way of homomorphisms between families of mental qualities and families of perceptible properties. The relations of similarity and difference that, e.g., mental shapes bear to each other parallel the relations of similarity and difference that perceptible shapes bear to each other. Accordingly, the mental quality of a visual sensation of a square, e.g., represents a perceptible square in virtue of resembling and differing from other mental shapes in ways parallel to the ways that perceptible squares resemble and differ from other perceptible shapes. Just as perceptible squares are more similar to perceptible rectangles than perceptible triangles, mental square, or square\*, is more similar to mental rectangular, or rectangular\*, than to mental triangular, or triangular\*.

Nevertheless, such mental qualities are not concepts, since concepts arguably do not represent properties and objects in virtue of homomorphisms between families of concepts and families of properties or objects. Concepts do not form families in the ways that mental qualities do. So mental qualities arguably represent perceptible properties nonconceptually.

Homomorphism theory also accounts for the additional aspects of perceptual experience that Peacocke (1983) discusses in his example of one's visual experience of the two trees, and it does so without committing to the existence of any nonrepresentational sensational properties of perceptual

experiences.<sup>18</sup>

As I argued above, objects have perceiver-relative properties, or P-properties. And two objects of the same size but at different distances from a perceiver have different P-properties in virtue of which they subtend different retinal angles; those objects have different P-sizes. And when one sees the two objects, one sees their P-sizes. As I argued above, seeing the size of an object and seeing how far away the object is located depends on one's seeing the P-size of the object. When one sees two objects of the same size but at different distances, one sees those objects as being the same size and as being at different distances away. But one also sees their different P-sizes. We can best explain how one sees the P-sizes, and other P-properties of objects, in terms of the view that one's sensations have mental qualities that represent those P-properties.

On this view, when one sees the two trees in Peacocke's example, one's visual sensation of the closer tree has a mental quality that pertains to that tree's P-size, and one's visual sensation of the farther tree has a mental quality that pertains to a different P-size. One's visual sensation of the closer tree is P-large\*, and one's visual sensation of the farther tree is P-small\*. Further, the mental quality of one's sensation pertaining to the P-size of the closer tree could

---

<sup>18</sup> I am of course not arguing that perceptual experiences have no nonrepresentational properties, just that we need not commit to their having nonrepresentational sensational properties to account for Peacocke's example. Perceptual experiences do of course have nonrepresentational properties. My current visual experience, e.g., has the nonrepresentational property of occurring at 5:30 pm.

resemble the mental quality of one's sensation pertaining to the P-size of a larger tree at a different distance more than it resembles the mental quality of one's sensation pertaining to the farther tree in Peacocke's example. Two trees of different sizes and distances sometimes look in some way the same size, or occupy the same amount of the visual field, because those trees cause sensations with the same mental qualities pertaining to P-size, i.e., they cause sensations with the same P-sizes\*.

Homomorphism theory explains how sensations represent the P-properties of objects. When one sees an object, one has a visual sensation with mental qualities that resemble and differ from other such mental qualities in ways parallel to the ways the P-properties of the object resemble and differ from other P-properties. For example, the P-size of a 10 ft. tall object 10 ft. away from one is more similar to the P-size of a 10 ft. tall object 9 ft. away than the P-size of a 10 ft. tall object 20 ft. away. Likewise, the visual sensation one has when seeing a 10 ft. tall object 10 ft. away has a mental P-size, or P-size\*, that resembles the P-size\* of a visual sensation of a 10 ft. tall object 9 ft. away more than it resembles the P-size\* of a visual sensation of a 10 ft. tall object 20 ft. away. The visual sensations one has when seeing the two trees in Peacocke's example have different P-sizes\*; one's visual sensation of the closer tree is P-larger\* than one's visual sensation of the farther tree.

Homomorphism theory also accounts for our sensing P-properties of objects other than their P-sizes. When one sees a circular object, such as a

penny, tilted slightly away from one's line of sight, one's visual sensation has a mental P-shape, or a P-shape\*. And that P-shape\* is the same as the P-shape of the visual sensation one has when seeing an elliptical object straight on; both sensations are elliptical\*. That is why we claim that circular objects tilted away from us look in some way elliptical, or at least that they look similar to elliptical objects seen straight on. Nevertheless, the visual system is sensitive to other factors in virtue of which it determines the difference in shape between an elliptical object seen straight on and a circular object tilted from one's line of sight. Because the visual system computes the perceiver-independent shape of the object, one sees the titled object as circular and the elliptical object as elliptical.

Homomorphism theory thus accounts for the qualitative character of perceiving the spatial properties of objects without committing to nonrepresentational, sensational properties of perception.

I have argued that we can account for the visual phenomenon Peacocke describes without committing to nonrepresentational, sensational properties of perceptual experiences. I argued that it could be that one's visual experience of the two trees in Peacocke's example represents the trees as being the same height while also representing them as having different perceiver-relative properties, P-sizes. Their representing the two trees as having those different P-sizes, I claimed, could account for the phenomenon by which the closer tree in some way occupies more of one's visual field than the farther tree. I further

argued that such representation is arguably nonconceptual, sensory representation, and I offered homomorphism theory as an account of such representation.

But Peacocke offers two other arguments for the claim that we must posit nonrepresentational, sensational properties to fully account for the qualitative character of perceptual experiences. And it could be that these two arguments succeed where the first fails. I will argue that they do not.

Peacocke claims there are cases in which the representational content and some nonrepresentational, sensational aspect of perceptual experience vary independently of each other. He claims that there are perceptual experiences within the same sensory modality that have the same representational content but differ in some other intrinsic, sensational respect, and that there are perceptual experiences that differ from each other in respect of their representational contents but that share some nonrepresentational, sensational aspect.

Peacocke cites cases of depth perception to argue that two perceptual experiences could have the same representational content while differing in some nonrepresentational, sensational aspect (1983, pp. 13-16). When one looks at an array of furniture in a room, one sees some of the items of furniture as being behind other items; one sees the pieces of furniture as being at different depths. And when one closes one eye and looks at the same room, one still sees those items of furniture as being at various depths. However, Peacocke

claims, the experience one has when looking at the room through one eye differs from the experience one has when looking at the room through both eyes. Since the two experiences both represent the pieces of furniture as being at various depths, he argues, the difference between those experiences is a difference in some nonrepresentational, sensational property associated with depth perception.

But one might argue that the difference between these two experiences is in fact representational. Perhaps binocular vision represents depth in some way that differs from the way monocular vision represents depth. Since binocular vision does in fact enable better depth perception than monocular vision, binocular and monocular experiences of depth do in fact differ in terms of the functional roles they play. Binocular experiences result in more accurate judgments of depth, and more accurate behavior directed towards objects. So binocular and monocular experiences of depth arguably differ in terms of how they represent depth. Perhaps that functional, representational difference between the experiences captures the difference Peacocke mentions.

However, Peacocke argues that if the difference between binocular and monocular experiences is purely representational, "... it ought to be impossible to conceive of cases in which the alleged sensational property is present, but in which a representation of certain objects as being behind others in the environment is absent" (1983, p. 14). Since, Peacocke further argues, we can in fact conceive of such cases, we cannot account for the difference between

monocular and binocular vision solely in terms of a difference in the representational properties of the experiences. Rather, such experiences also differ in respect of some nonrepresentational, sensational properties.

Peacocke's reasoning is as follows. If the difference between the two visual experiences of depth is a purely representational difference, e.g., a difference in the precision or accuracy with which they represent depth, then we could not imagine a situation in which one has an experience that does not represent depth at all but that does have the aspect present in the binocular case but not in the monocular case. If that difference is purely representational, then when one imagines a case in which one's experience does not represent depth at all, one would imagine an experience that lacks the aspect that is present in the binocular but not the monocular experience.

Peacocke claims that we can imagine perceptual experiences of depth that nonetheless do not represent anything as being at any depth. And he further argues that we could imagine a case in which one has a visual experience of depth that has the aspect that differentiates binocular and monocular experiences of depth, but that does not represent anything at a depth.

Peacocke's example invokes the perceptual experiences generated by prosthetic vision. Paul Bach-y-Rita (1972) developed a prosthetic device that substitutes tactile stimulation for visual stimulation. The system translates information from a video camera mounted on a pair of eye-glass frames into vibrations on a matrix of pins placed against some patch of skin on a blind

person's body, e.g., on the person's back or tongue. When properly trained, people can use this tactile-visual substitution system (TVSS) to navigate their environments and even to identify objects with some success. The system provides the user with spatial information about objects and the layout of the region of space in front of the camera. However, before subjects are fully trained on the TVSS, there is a period during which they claim that the sensations they have as a result of the TVSS are neither sensations of anything as existing out in the space in front of them, nor tactile sensations of a stimulus on their skin. Still, the sensations vary in two dimensions, corresponding to the two spatial dimensions of the vibrotactile array causing them; in this sense, they are spatial experiences.

Peacocke supposes that we can imagine adding a third dimension of variation to those sensations, e.g., by adding a second camera, thus creating a binocular TVSS. In this case, Peacocke claims, the TVSS user would have sensations of three dimensions of space that nonetheless do not represent anything as being at any depth. This experience, Peacocke asserts, would be an experience of depth that does not represent depth. And, he further claims, we could imagine that a blind person who is suddenly given binocular vision would have a visual experience of depth that does not in fact represent anything as being at any depth, just as the binocular-TVSS user has a TVSS experience of depth that does not represent anything as being at any depth. When we imagine such an experience, Peacocke assumes, we imagine an experience with the

nonrepresentational aspect of seeing depth that is present in normal binocular visual experiences of depth but not in normal monocular visual experiences of depth. So, he concludes, visual experience involves such a nonrepresentational aspect.

Peacocke's argument rests on the assumption that there would be a stage during which the binocular-TVSS user has sensations that do not represent depth. And Peacocke assumes this because, at some point during their training, TVSS users have sensations that "...do not seem to [them] to be of objects in the space around [them]... The subjects report that the sensations are not as of anything 'out there'" (1983, p. 15).

But we can account for the occurrence of sensations of depth that one nonetheless claims are not sensations of anything at any depth in terms of the view that sensations are nonconceptual representations. The binocular-TVSS user in Peacocke's imaginary case has an experience with three dimensions of variation. But that being, we imagine, reports the experience as not being of anything at any depth. Nevertheless, even if the being reports the experience as not being of anything at any depth, the experience could in fact represent something at a depth. One's reports express one's intentional states, such as one's thoughts and beliefs; intentional states are prerequisites for reporting. So one would report that one sees something at a depth, or as being out in the space surrounding one, only if one thought that there was in fact something at a depth. But having a nonconceptual, sensory representation without having such

a thought is insufficient for such a report. If one had a nonconceptual, sensory representation of something at a depth, but one did not have an accompanying intentional, conceptual representation of something at a depth, then one would deny seeing something at a depth while having that three-dimensional experience. And it could be that the binocular-TVSS user Peacocke imagines would not form the intentional representation, e.g., the belief, that there is something at a depth because he or she has not yet forged a connection between such intentional states and the novel nonconceptual, sensory representations caused by the binocular TVSS. Peacocke's example thus does not show that binocular experiences of depth perception have some nonrepresentational, sensational property that monocular experiences of depth do not have.

Peacocke also claims that there are visual experiences that differ in their representational content, but that have some nonrepresentational qualitative aspect in common (1983, pp. 16-17). For example, imagine looking through one eye at a wire figure in the shape of a cube, where side ABCD of the cube is in front of side EFGH. One can at one moment see ABCD as in front of EFGH, and at the next moment see ABCD as being behind EFGH. This aspectual switch in how one sees the cube reflects a switch in the representational content of one's experience of the cube. However, Peacocke claims, there is also some aspect of one's experience of the cube that remains constant between these switches in

representational content. That aspect, Peacocke claims, is a nonrepresentational, sensational property.

Peacocke claims there is a common aspect between the experiences because one sees that the cube has not changed, even though one sees it differently in each case (1983, p. 16). But Peacocke does not explain why one's seeing the cube as invariant is not a function of the representational content of that experience. One's seeing that the cube isn't changing is arguably a function of one's beliefs about the cube's not changing. And such beliefs are paradigmatic representational states.

Perhaps Peacocke would reply that we can best explain why one forms that belief on the basis of some nonrepresentational feature that is common to both experiences. But one could form such a belief on the basis of factors other than a common, nonrepresentational feature of those experiences. For example, it could be that one forms that belief because one believes that the sides of objects, such as the cube, do not reverse themselves without those objects' moving, and because one did not see the cube move.

In any case, perhaps this example presents a difficulty similar to that posed by Peacocke's example of the two trees. If at one moment one's visual experience represents face ABCD of the cube as being in front of face EFGH, and at the next moment one's experience represents ABCD as being behind EFGH, but one's experience also represents the cube as not changing, then one's experience represents a contradiction. Since an experience cannot

represent a contradiction, Peacocke could argue, one's experience does not represent ABCD as being in front of EFGH at one moment, ABCD as being behind EFGH at the next moment, and the cube as not changing throughout.

However, even if one's initial experience of the cube represents ABCD as being in front of EFGH, and one's subsequent experience of the cube represents ABCD as being behind EFGH, it could be that one does not form both the belief that ABCD is in front of EFGH and the belief that ABCD is in back of EFGH. Rather, it could be that one forms the belief that ABCD is in front of EFGH, and then one has a nonconceptual visual representation of ABCD behind EFGH without forming the belief that ABCD is in fact behind EFGH. If so, then one's representations of the relative positions of ABCD and EFGH do not contradict each other, and they do not contradict one's belief that the square does not change while one is performing the experiential switch.

As with Peacocke's other examples, we need not posit nonrepresentational, sensational properties to account for this example.

## Chapter 2: Change Blindness, Part 1

### 1. Introduction

In the previous chapter, I argued that we sense the spatial properties of stimuli in virtue of having sensory states, i.e., sensations, that have properties pertaining in a specific way to those perceptible spatial properties. Such properties of sensations, I argue, are distinctly mental, and they correspond to the spatial properties of stimuli by way of homomorphisms between families of perceptible spatial properties and families of mental qualities. Accordingly, a visual sensation of a square, e.g., has a mental quality that occupies the same position in its property family that perceptible squareness occupies in the family of perceptible shapes; just as physical, perceptible squares resemble physical, perceptible rectangles more than physical, perceptible triangles, the property of a visual sensation of a square, i.e., square\*, resembles rectangular\* more than triangular\*. This homomorphism theory, I argued, can account for our sensing not only shapes but also all other perceptible spatial properties, such as the locations, sizes, orientations, and movements of objects, as well as their nonspatial properties, such as color, texture, and temperature.

This view suggests that one sees the spatial layout of a visual scene in virtue of having a sensation that has mental qualities representing that spatial layout.

But recent experiments on the phenomenon of change blindness intuitively seem to challenge the view that one sees the spatial layout of a visual scene in virtue of having a sensation with properties that pertain to that spatial layout. Subjects in change-blindness experiments often fail to notice significant changes in visual scenes. For example, one could fail to notice that a central figure of a scene is changing locations, size, or color when that change occurs while the scene is obscured, or while one is looking elsewhere. When subjects in such experiments are subsequently told what in the scene changed, they immediately notice it and express great surprise at their having missed it.

Intuitively, such change blindness poses a challenge to the view that one sees the spatial layout of a scene in virtue of having a visual sensation with mental qualities pertaining to the spatial properties of the scene. If one had such sensations with such mental qualities, the sensation one had before the change and the sensation one had after the change would presumably differ in ways corresponding to the changed perceptible features. And, one might argue, if one's sensations changed, one would presumably notice that change, so one would notice the change in the visual scene. Since one fails to notice changes in visual scenes during change-blindness experiments, one might further argue, one does not have such sensations with mental qualities pertaining to the spatial layout of the visual scene. Alternatively, if we do have such sensations with mental qualities pertaining to the spatial layout of visual scenes, as I have

argued, we must explain why we sometimes fail to notice changes in those visual scenes.

In this chapter, I'll examine the psychological literature on change blindness and several accounts of what change blindness reveals about visual perception. I'll focus primarily on Alva Noë's view that change blindness supports his so-called enactive theory of visual perception against the orthodox view that visual perception involves visual representations, e.g., sensations, of the spatial layout of visual scenes. I'll argue that change blindness does not in fact support Noë's view, nor does it challenge the view that we see the spatial layout of visual scenes in virtue of having sensations with mental qualities pertaining to the spatial properties of those scenes. In so doing, I'll argue that we can best explain visual perception in terms of the existence of visual representations, such as sensations, not in terms of Noë's view that visual perception does not rest on such representations.

Before examining Noë's account of change blindness and visual perception, I will first discuss a number of experiments on the phenomenon of change blindness.

## 2. Experiments on Change Blindness

Change blindness occurs under a number of different conditions, including both highly controlled experimental settings and real-life situations.

Some of the earliest experiments on change blindness tested one's ability to see changes in visual stimuli when those changes occurred during saccades, quick movements of one's eyes that occur three or four times a second. One usually fails to notice these eye movements, which often occur involuntarily. However, they play a significant role in perception, enabling one to fixate a number of stimuli in a short period of time. The early experiments on change blindness showed that subjects often fail to notice changes to visual scenes when those changes occur during subjects' saccades.

The earliest of these studies examined how changing the visual properties of text affects one's ability to read that text. George McConkie and David Zola (1979) tested how altering the cases of letters affects one's ability to read a sentence. To do so, McConkie and Zola used the so-called eye-movement-contingent display system, developed by McConkie and Keith Rayner (see McConkie, Zola, Wolverton, and Burns, 1978), in which an eye-tracking device monitors one's saccades and triggers a computer to change a visual scene during those saccades.

McConkie and Zola presented subjects with a sentence printed in alternating capital and lowercase letters, e.g., sentence (1) below.<sup>19</sup>

1) ThE sPaCe ShUtTIE tHuNdErEd InTo ThE sKy On A cOIUmN oF sMoKe.

---

<sup>19</sup> Since reading such sentences is difficult, subjects were first habituated to a number of similar sentences, also composed of alternative capital and lowercase letters.

When subjects saccaded, the computer switched the case of each letter, changing capital letters to lowercase letters and lowercase letters to capital letters. Sentence (1), e.g., was changed to sentence (2) below.

2) tHe SpAcE sHuTtLe ThUnDeReD iNtO tHe SkY oN a CoLuMn Of SmOkE.

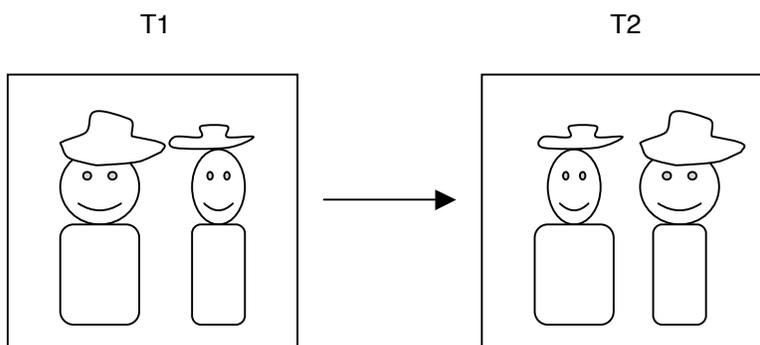
The computer alternated between such sentences each time the subject saccaded.

McConkie and Zola found that subjects did not notice that the letters were changing case. Also, before running the experiment on subjects, they ran it on Zola, who expressed concern that the setup was malfunctioning, and that the computer was not switching the cases of the letters at all. However, though Zola failed to notice the changes, the experimenters whose eye movements were not being monitored by the eye tracker easily noticed the changes as they watched the computer screen.

John Grimes (1996) used an eye-movement-contingent display system to test subjects' ability to detect changes in photographs when those changes are made during subjects' saccades. Subjects were presented with a photograph, e.g., of two cowboys sitting on a fence, or of a city's skyline. Like in McConkie and Zola's experiment, subjects' eye movements were monitored by an eye-tracking device. As soon as a subject saccaded, the computer switched the

photograph with another photograph that differed from the first in respect of one prominent detail. For example, in the trials in which subjects were presented with a photograph of two cowboys, the heads of the two cowboys were switched in the second photograph. And in trials in which subjects were presented with a photograph of a skyline, a prominent skyscraper was 25% larger in the second photograph than in the first photograph.

Figure 1: Adapted from Grimes (1996)



Grimes presented subjects with a photograph, e.g., of two cowboys. When subjects saccaded, something in the photograph changed, e.g., the heads of the two cowboys switched places.

Grimes found that none of the subjects shown the photographs of the skyline noticed the change in the size of the skyscraper, and only 50% of the subjects presented with the photographs of the cowboys noticed that their heads were being switched. Grimes tested subjects with 10 different pairs of photographs, and found those subjects noticed only 33% of the changes made to those photographs.

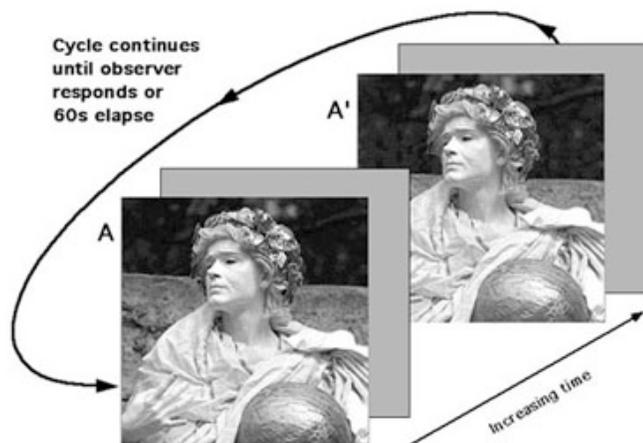
On one widely held view, one fails to notice changes to a visual scene made while one is saccading because one normally sees such changes by detecting what are called motion transients, and one's saccades prevent one from seeing those motion transients. Motion transients are slight flickers caused by a changing feature that serve as a signal, stimulating visual processing of features at that location. Saccades produce motion transients of their own, since all retinal stimulation changes when one's eyes move. And such changes are much like the changes caused by a change at a single location in one's field of view. So saccades produce global motion transients, corresponding to changes at every location in one's field of view. And it could be that such global motion transients mask the local motion transients caused by the changes in the sentences in McConkie and Zola's experiments and the changes in the photographs in Grimes's experiments.

To test whether change blindness is in fact caused by one's failure to detect local motion transients, and not by some other factor particular to saccading, Ronald Rensink, Kevin O'Regan, and James Clark (1997) developed the so-called flicker paradigm. The flicker paradigm controls for the global motion transients produced by saccades by enacting changes in visual scenes independently of subjects' saccades. Subjects are presented with a picture that briefly disappears at regular intervals. After those brief interruptions in the visual scene, the picture reappears. However, when it reappears, the picture is changed in respect of some significant feature. Then the picture disappears

again for another brief interval, after which it reappears in its original form, and the cycle begins again until the subject finally notices the change or the experimenter ends the trial.

The pictures used in such flicker experiments, like those used in Grimes' experiments, involve significant changes. For example, a subject is presented with a picture that changes between an image of the Cathedral of Notre Dame in Paris and an image of the same cathedral missing one of its two towers, or the subject is presented with a picture of an airplane with a jet engine attached to its wing alternating with a picture of that plane missing its jet engine.

Figure 2: Flicker Paradigm<sup>20</sup>



Subjects are briefly presented with picture A, followed by an intermittent blank screen, then by picture A', which differs from A in some respect, and then by another blank screen. The cycle then begins again. In this example, the wall behind the statue is higher in A than in A'.

<sup>20</sup> From Ron Rensink's website, <http://www.psych.ubc.ca/~rensink/flicker/>. Image reprinted with permission from Ron Rensink. For demonstrations of the flicker paradigm, see Dan Simons's website: <http://viscog.beckman.uiuc.edu/change/demolinks.shtml>.

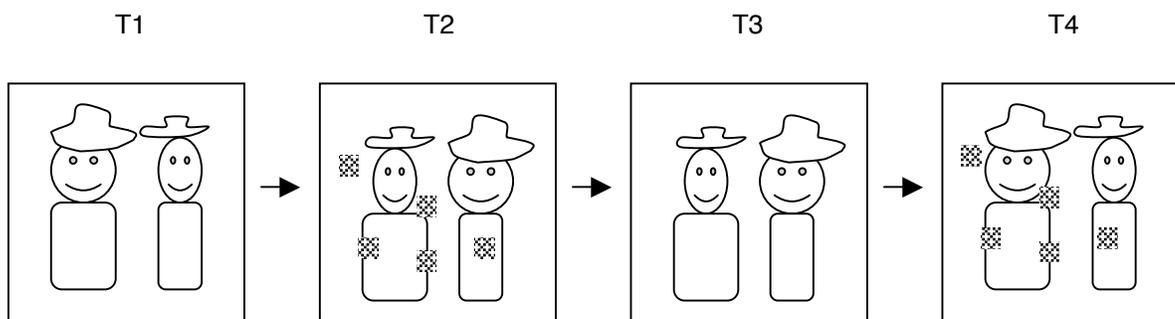
Rensink et al. found that subjects often fail to notice these changes during the first cycle of the trial, and many fail to notice them even after a minute of cycles. However, when one sees these pictures alternating without the intermittent blank screens, one immediately notices the changes. And when subjects finally discover the changes, or the experimenter describes the changes to them, the subjects are often greatly surprised that they missed the changes. These results were also found in experiments in which changes to scenes were made while subjects blinked (O'Regan, Deubel, Clark, and Rensink, 2000).

The flicker-paradigm experiments show that subjects fail to notice significant changes in pictures even when those changes are made independently of subjects' saccades. So these experiments show that change blindness does not result from something specific to one's saccades. However, since the changes are made to the pictures during the intermittent blank screens, so when subjects do not see the pictures, those changes occur without subjects' seeing the motion transients normally caused by such changes. So perhaps change blindness does in fact result from one's failure to see local motion transients caused by local changes to scenes.

However, experiments using another paradigm challenge this view. O'Regan, Rensink, and Clark (1996) showed that subjects often fail to notice changes in pictures even when the motion transients produced by those changes are visible. Motion transients due to changes made during saccades, during an intermittent blank screen, or during a blink are not visible; they are masked by the

global motion transients caused by one's eye movements, they are obscured by one's eyelids during a blink, or they are never produced by the change, as in the flicker paradigm. To test whether change blindness results from one's failing to see local motion transients, O'Regan et al. developed the so-called mudsplash paradigm, in which a picture is changed while an unaltered part of the same picture is briefly occluded by colored shapes simulating a splash of mud on the windshield of a car.

**Figure 3:** Mudsplash Paradigm



Subjects are shown a picture at T1. At T2 a simulated mudsplash appears and the scene changes; e.g., the heads of the two figures switch places. At T3 the simulated mudsplash disappears. At T4 the mudsplash reappears, and the scene changes again, i.e., the heads switch places again. The cycle then begins again.

O'Regan et al. found that subjects fail to notice significant changes in pictures when those changes occur during such simulated mudsplashes, even though the changes are not themselves obscured by the simulated mudsplash.

Again, subjects easily notice those changes when they occur in the absence of a simulated mudsplash. Since the changes are not obscured in these cases, the motion transients they produce are not obscured either. So change blindness can occur even when such motion transients are visible.<sup>21</sup>

However, even though the change in the scene is not obscured in mudsplash experiments, the appearances and disappearances of the mudsplashes themselves also present changes in the visual scenes. So it could be that the motion transients caused by the mudsplashes mask the motion transients caused by the change in the visual scene.

But change blindness can also occur in the absence of any disruption at all in the visual scene, e.g., a disruption due to a saccade, blink, or intermittent blank screen, or any distractor, such as a mudsplash. Dan Simons, Steven Franconeri, and Rebecca Reimer (2000) and Cédric Laloyaux, Christel Devue, Elodie David, and Axel Cleeremans (submitted) showed that subjects are even worse at noticing significant changes in scenes when those changes occur gradually in front of their open, functioning eyes than they are at noticing changes that occur during flicker-paradigm experiments. In such gradual-change experiments, subjects are presented with a scene that changes slowly over the

---

<sup>21</sup> Perhaps there is some other reason to think that the mudsplashes render the local motion transients caused by the change invisible. For example, one might argue that the mudsplash draws one's attention away from the change, and thus from the location of the motion transient. If so, and if attention is necessary for seeing, one would fail to see the local motion transient caused by the change. I'll discuss the view that change blindness results from a failure to attend to a changing feature below.

course of 12 seconds. For example, subjects are presented with a picture of a house with a chimney. Over the course of 12 seconds, the chimney gradually fades into the background, leaving a picture of the house without a chimney.<sup>22</sup> At the end of the trial, subjects were prompted to use a mouse to click on the region of the picture that changed. They were also asked to report whether they saw a change and, if so, whether they were confident that they saw a change where they clicked, whether they simply guessed at the location of the change, or whether they thought they saw a change but were not sure that they did. Both Simons et al. and Laloyaux et al. found that subjects failed to notice a significant number of such gradual changes to pictures.

Simons et al. also ran a flicker experiment using the image from the beginning of the gradual-change trials and the images from the end of those trials. In this case, the picture of the house with the chimney was presented for 11,250 msecs, then a blank screen was presented for 250 msec, and finally the picture of the house without the chimney was presented for 11,250 msecs. Simons et al. found that subjects in this flicker experiment were slightly more successful at noticing changes than subjects in the trials in which the picture changed gradually, though they still failed to notice a significant number of changes.

---

<sup>22</sup> In half of the trials, the change was reversed. In this case, such a reversal consisted of a picture of a house without a chimney gradually morphing into a picture of a house with a chimney.

In the gradual-change experiments, as opposed to the flicker experiments, the change in the picture occurs in front of one's open, functioning eyes, and, unlike in the saccade experiments and the mudsplash experiments, no global or local motion transients occur that could mask the gradual change.<sup>23</sup> Change blindness can thus occur even when the change in the visual scene occurs in front of one's open, functioning eyes and in the absence of a masking change.

Massimo Turatto, Alessandro Angrilli, Veronica Mazza, Carlo Umiltà, and Jon Driver (2002) found that change blindness occurs more often with changes in the background of a scene than it occurs with changes in the foreground, even if those background changes are more significant than the foreground changes. They also found that semantic primes reduce subjects' rates of change blindness.

Turatto et al. presented subjects with a scene consisting of six dots, some of which were light gray and some of which were dark gray, arranged in a circle at the center of the screen against a background consisting of 20 alternating black and white stripes. The initial scene appeared for 400 msec and was followed by a blank screen for 100 msec. Then a second scene also consisting of six gray dots arranged in a circle against a background of black and white vertical stripes appeared for 400 msec.

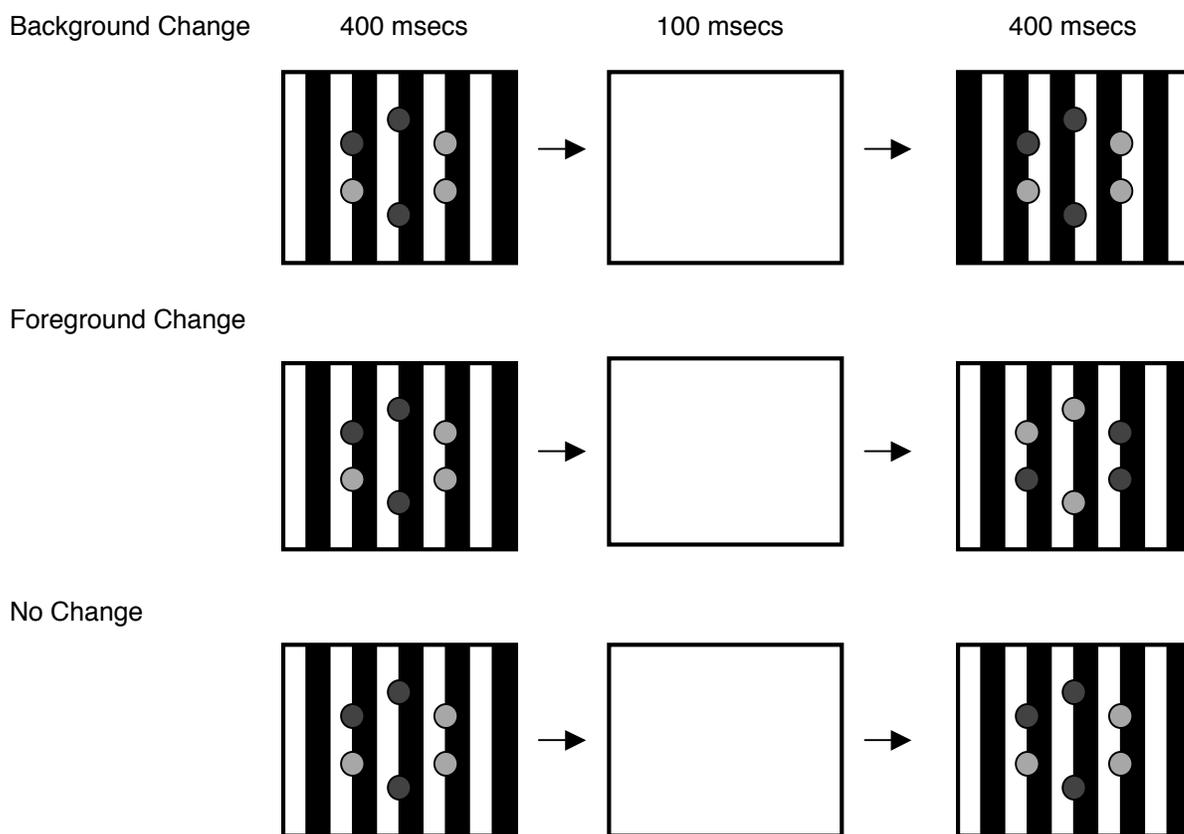
---

<sup>23</sup> Both Simons et al. and Laloyaux et al. ran similar experiments in which an object in one picture changed color. And Laloyaux et al. ran experiments in which the facial expressions of people in pictures changed. The results of all of these experiments showed significant change blindness in both the gradual and flicker trials.

The scene could either undergo a change after the blank screen or it could remain the same. A foreground change consisted in each of the six dots changing luminance, with light gray dots changing to dark gray dots and dark gray dots changing to light gray dots. And a background change consisted in each of the vertical stripes changing color, with black stripes changing to white stripes and white stripes changing to black stripes.

After the second scene disappeared, subjects were to report whether they saw any change at all.

Figure 4: Adapted from Turatto et al. (2002)



The first set of trials consisted in non-cued trials in which subjects were not primed to look for a particular kind of change. Each trial began after the subject heard the word 'attention' for 500 msec. In these trials, subjects reported only 10% of the changes that occurred to the background. However, this is not significantly different from the 13% rate of false alarms during trials in which no change occurred. Further, all of the subjects in these trials expressed surprise when they were told about the background changes they had missed. However, subjects correctly reported 98% of the changes to the foreground dots.

Turatto et al. also found that semantic priming significantly reduces change blindness. They ran a second block of trials in which subjects were cued to look for a particular kind of change. Some trials began after the subject heard the word 'background', indicating that if a change occurred, it would occur in the background. And other trials began with the word 'circles', indicating that if any change occurred, it would occur in the foreground circles.

In these cued trials, subjects correctly reported 88% of the background changes, suggesting that the questioning at the end of the first block of trials and the word 'background' cued them to look for the changes in the background stimuli. This suggests that change blindness is reduced by semantic priming, which in turn suggests that the mechanisms that enable one to successfully

notice changes, i.e., the mechanism that is inoperative during change blindness, is susceptible to semantic priming.<sup>24</sup>

All of the change-blindness experiments I have discussed so far involve cases in which one fails to notice changes in static images, such as photographs, static computer images, or text. However, change blindness also occurs with dynamic images, such as movies, and it also occurs in live situations. This suggests that change blindness is not limited to picture perception, and it is not an artifact of experimental settings.

Daniel Levin and Daniel Simons (1997) tested whether subjects would notice significant changes in movie scenes when those changes occurred across edits, e.g., when they occur during cuts from one camera to another. In the first experiment, Levin and Simons presented subjects with a short film of two people sitting at a table and talking. There were a total of nine film cuts during the movie, and some element of the scene changed during each of these cuts. For instance, one of the characters was wearing a scarf that disappeared after one cut, and plates resting on the table in front of the people changed color during another cut. Levin and Simons found that only one out of 90 subjects noticed any of the changes occurring during the film.

---

<sup>24</sup> Other experiments showed that subjects' familiarity with the subject matter of changing pictures also effects their rates of change blindness. For example, drug users notice more changes in pictures of drug paraphernalia than people who do not use drugs (Jones et al., 2003). And football experts notice more changes in football scenes than non-experts (Werner and Thies, 2000).

These changes made during the film were somewhat peripheral and arguably unimportant to the story line. However, in a second experiment, Levin and Simons showed subjects a film with a single character, in which the actor playing that character changes during an edit. In one such film, a person sitting at a desk hears a phone ring and gets up to go and answer the phone. When the person stands up, the film cuts to a different camera angle, and the actor who appears after that cut is different from the actor who appeared before the cut.

Subjects in these experiments were given no directions before viewing the film, and they were asked to write a description of it afterwards. If a subject failed to mention the change in actor, the experimenters asked the subject whether he or she noticed the switch. Levin and Simons found that only 33% of the subjects noticed that the actor was switched during the film.<sup>25</sup> However, even subjects who failed to notice the change provided otherwise detailed descriptions of the films.

Simons and Levin (1998) also showed that change blindness occurs in real-life situations. In their experiment, an experimenter stopped students on a college campus to ask for directions. While the experimenter and the student

---

<sup>25</sup> The pairs of actors used in these films were of the same gender and race, had the same hair color, either both wore glasses or neither did, and they wore similar clothing. However, in a subsequent experiment, subjects were shown both films in which the actors were switched and films in which they were not switched, and they were instructed to look for changes occurring during the film. Those subjects had little trouble identifying the changes in actors. Since the same actors were used in these films, this suggests that subjects' failure to notice the switch in the previous experiment was not due to the similarities between the actors.

were talking, two other people carrying a door rudely walked between the experimenter and the student, interrupting their conversation. As the people carrying the door passed between the experimenter and subject, the experimenter changed places with one of them, and that person then stayed behind and continued the conversation with the student.

Simons and Levin found that only half of the subjects in this experiment noticed that the person they were speaking with after the interruption was different from the person they were speaking with before the interruption. These results suggest that people fail to notice significant changes in visual scenes even in everyday life, i.e., that change blindness is not an artifact of psychological experiments run in laboratories, nor is it an artifact of picture or movie perception.

There are at least two crucial findings of these experiments on change blindness. One is that people fail to notice significant changes in visual scenes, and that such failures to notice changes occur under a variety of different circumstances. Another interesting finding is that subjects are greatly surprised that they fail to notice many of these changes. The first finding, one might argue, seems to suggest something about visual perception, e.g., that we fail to see significant features in visual scenes. The second finding, one might argue, reflects a folk psychological commitment to the view that we see a great amount of detail in visual scenes, and perhaps that our visual experiences of visual

scenes are themselves highly detailed. Any theory of visual perception must account for these two findings.

### 3. Change Blindness and Visual Representations

Perhaps the most radical account of change blindness is the view that it shows visual perception involves no visual representations of the spatial layouts of visual scenes. If visual perception involved such representations, such as sensations with mental qualities pertaining to the spatial properties of stimuli in one's field of view, then one's representation would have properties pertaining to the spatial layout of the visual scene one is currently viewing. However, one might argue, if one sees a visual scene in virtue of having a sensation with mental qualities pertaining to the spatial layout of that scene, it is unclear why one would fail to notice an otherwise obvious change in that scene, e.g., a change in the location, size, or color of a significant figure.

Alva Noë (2004, 2005; O'Regan and Noë, 2001) claims that change blindness poses such a challenge to the view that we see visual scenes in virtue of having visual representations, such as sensations, that represent the spatial layouts of those scenes. In particular, Noë argues that change blindness raises a problem for what he calls the orthodox view of visual perception, according to which visual perception involves the construction of highly detailed, picture-like representations of visual scenes, i.e., representations with mental qualities representing the spatial properties of those scenes. This view, which he

attributes to Ernst Mach (1906/1959), is motivated primarily by phenomenological concerns. According to the orthodox view, we can best explain why visual perception seems to present one with so much detail of visual scenes in terms of the view that one sees such scenes by having detailed visual representations of them. If visual perception does in fact involve such detailed representations, it is unclear why change blindness occurs. If one sees a scene in virtue of having detailed representations of the features of that scene, then one will presumably have detailed representations that represent the features of the different scenes presented in succession in change-blindness experiments. And, if the visual representations representing the features of different, consecutively presented scenes differ in ways pertaining to the differences between those visual scenes, one would presumably notice the changes in those scenes. Since subjects in change-blindness experiments often fail to notice such changes, Noë argues, visual perception must not involve such detailed representations.

To be more precise, Noë does not claim that change blindness shows that there are no detailed visual representations of visual scenes. He claims that “[c]hange blindness is compatible with the existence of detailed internally stored information about what is present to vision” (2004, p. 52). Rather, according to Noë, “[c]hange blindness suggests that we don’t make use of detailed internal models of the scene (even if it doesn’t show that there are no detailed internal representations). In normal perception it seems that we don’t have online access to detailed internal representations of the scene” (2004, p. 52). So Noë holds

that if one made use of detailed visual representations, or if one had online access to them, one would not fail to notice significant changes in visual scenes, as one often does in change-blindness experiments.

Noë is presumably equating one's using a visual representation and one's having online access to a visual representation. And one's having online access to a visual representation is arguably a matter of one's being conscious of that visual representation. But one could arguably use a visual representation without being conscious of the representation, i.e., if the representation mediates visual stimulation, other psychological states, and behavioral outputs while one is unaware of that visual representation. So it could be that one uses visual representations in this way, even when one fails to notice changes in visual scenes. I'll address this issue in the next chapter, and I'll argue that we do in fact use visual representations in this way, even when we fail to notice changes in visual scenes. I'll argue that this shows that change blindness is a failure to be conscious of seeing a change, not a failure to see a change.

However, in this chapter I'll address Noë's claim that change blindness shows that we do not use visual representations, where such use involves one's conscious access to those representations. I'll argue that change blindness does not threaten the view that visual perception involves visual representations, though perhaps it suggests that those representations are not always highly detailed.

Noë claims change blindness shows that one sees very little of the visual scene in front of one's eyes at any given moment. According to Noë, one sees only what one attends to. "If a change takes place when attention is directed elsewhere, the change will tend to go unnoticed. In general, you only see that to which you attend. If something occurs outside the scope of attention, even if it's perfectly visible (i.e., unobstructed, central, large), you won't see it" (2004, p. 52).

However, Noë does not explain what he takes attention to be. So it is not clear how to understand his claim that we see only what we attend to.

Presumably, Noë assumes that attention requires no explanation; one simply knows what it is to attend to something.

Common sense arguably holds that visual attention is a process whereby vision allocates more resources to processing certain select stimuli than it allocates to processing other stimuli. Accordingly, attention is a limited-capacity mechanism that facilitates and heightens perception of some stimuli at the expense of other stimuli.<sup>26</sup>

But, one might argue, Noë's view that we see only what we attend to conflicts with such commonsense theorizing about visual perception. Though we

---

<sup>26</sup> This view of attention is widely held throughout psychology as well. But it is widely debated what kind of processing attention facilitates. In chapter 4, I discuss the views that attention facilitates our perception of combinations of distinct features, for example when seeing combinations of color, shape, size, and location, and that attention facilitates our ability to keep track of objects. Other views hold that attention facilitates working memory, our conscious awareness of our perceptions (Prinz, 2005), and our knowledge of what our words refer to (Campbell, 2002).

often take ourselves to see the objects we attend to better than the objects we do not attend to, i.e., that attention somehow heightens and facilitates our perceptual awareness of select stimuli, we do take ourselves to see objects that we fail to attend to. Attention does not simply block out everything one is not attending to. Also, objects one is not currently attending to often capture one's attention, e.g., when a flash of light draws one's attention from what one is currently attending to towards that flash of light. But it is unclear how the flash of light could capture one's attention if one did not already see the region of space where the flash of light occurred.<sup>27</sup>

Nevertheless, perhaps Noë's claim that change blindness undermines the so-called orthodox view does not rest on his assumptions about the relations between seeing and attending. Change blindness, Noë argues, shows that we do not see a great amount of detail at once, so vision does not involve detailed representations of visual scenes. And that could be true even if we do in fact see more than we visually attend to. If so, perhaps we see less than the orthodox view holds, but we see more than just those stimuli we attend to.

But, one might argue, this account of seeing conflicts with what it's like for one to see a visual scene. When one sees a visual scene, one seems to oneself to see a great amount of detail at once. This is, of course, the motivation for the

---

<sup>27</sup> The so-called cocktail-party effect presents an auditory analogue. When one is involved in a conversation at a cocktail, one attends to that conversation. However, when someone involved in another conversation in the room says one's name, it often catches one's attention. But it is not clear how one's name could catch one's attention if one did not hear it in the first place.

so-called orthodox view that Noë rejects. Again, according to the orthodox view, we seem to see a great amount of detail at any given moment because we do see a great amount of detail at any given moment; and we see a great amount of detail at each moment because vision constructs detailed representations of all the visible detail in the current visual scene. If Noë is right that we do not see a great amount of detail at once, he must explain why one seems to oneself to see a great amount of detail at once, and why one's visual experiences themselves seem to one to be so detailed.

Some (e.g., Blackmore et al., 1995; Dennett, 1991, 2002; O'Regan, 1992) argue that our impression that we see a great amount of detail at once is illusory. According to this view, we do not see a great amount of detail at once, and visual experience is not itself highly detailed, though it seems to us that we see a great amount of detail at once, and that our visual experiences represent a great amount of detail at once.

But Noë claims there is no such illusion in ordinary visual perception. Rather, he argues, it does not seem to one that one sees a great amount of detail at once, and one's visual experiences do not seem to one to be highly detailed. Rather, according to Noë, we take visual experience to present us with a great amount of detail because we can easily move in ways that will enable us to see details that we do not currently see, and we implicitly understand how such movements will enable us to see those details (2004, p. 63). In this way, one is aware of the visual scene as being highly detailed, and one is aware that

one can easily access those details. But, on Noë's view, visual experience presents a visual scene as detailed without presenting one with all of its details at once.

One might argue that this view fails to explain an important aspect of change blindness. When change-blindness subjects finally discover the change they previously failed to notice, or when the experimenter tells them what had been changing, subjects are greatly surprised that they failed to notice the change right away. That subjects are so surprised is one of the main motivations for the claim that we are under the illusion that visual experience is highly detailed and presents one with a great amount of detail at once. Since Noë claims that we are not under such an illusion about our visual experiences, he must explain why change-blindness subjects are so surprised when they discover the changes they were missing.

Noë claims that subjects are surprised, not because they think their experiences are so detailed, but because they think they are better at noticing changes than they in fact are (2004, p. 58). So, on Noë's view, subjects' surprise about their own change blindness reveals that they overestimate the ease with which they can access details in visual scenes, not that they overestimate the detailed nature of visual representations involved in visual experience.

Noë motivates his so-called enactive approach to visual perception, i.e., the view that we see visual scenes as detailed in virtue of implicitly understanding how one's movements will enable one to see more detail than one

currently sees, with a number of examples (2004, p. 60). In one such example, Noë draws an analogy between visual perception and tactile perception. When one grasps an object, e.g., a bottle, in one's hand, one touches only those parts of the object currently in contact with one's skin. Nevertheless, one does not feel the bottle as consisting of only those parts. Rather, one feels the bottle as having parts that one is not touching but that one could touch if one moved one's hands and fingers over the surface of the bottle. According to Noë, the impression that the bottle consists of more than those parts one is currently touching, results from one's understanding that if one moved one's hands in certain ways, one would feel other parts of the bottle.

Vision, according Noë, is analogous to touch in this way. Just as one feels a bottle one grasps as consisting of more than just the parts one currently touches, one experiences a visual scene as consisting of more than just those parts that one is in fact seeing at that moment. And just as one perceives the bottle as consisting of more parts than just those one is touching because one understands that one will feel other parts of the bottle if one moves one's hands in certain ways, one perceives the visual scene as consisting in more than just what one sees at that moment because one understands that if one moves one's eyes and head, one will see more than what one currently sees.

Noë illustrates his point with a visual example as well. When one sees a cat standing behind a picket fence, one sees only the parts of the cat that show through the fence, since the fence occludes the other parts of the cat.

Nevertheless, one sees the cat as a whole animal, not as one consisting of only those parts that one currently sees. According to Noë, this is because one understands how one could move in order to see the parts of the cat that are currently hidden from view; e.g., one understands that one could walk to the other side of the fence, or one could peer over the top of the fence.

Noë claims that all visual perception involves one's understanding that one could move in certain ways to see details in the visual scene that one does not see at that moment. Accordingly, we take our visual experiences to present us with highly detailed visual scenes, despite our seeing only a very limited amount of detail at each moment, because we understand that we can easily move to see details of the visual scene that we do not currently see, not because seeing involves the construction of detailed visual representations.

#### 4. Sparse Visual Representations

The view that we see very little at once is of course compatible with the view that we see stimuli in virtue of having visual representations of those stimuli and their properties. It could be that one sees only a small subset of the visible details of a visual scene at once, as Noë claims, but one sees those details in virtue of having visual representations, or sensations, that represent those few details (O'Regan, 1992; Simons & Levin, 1997). This would be the case if at any given moment one has a visual sensation with mental qualities pertaining to only a small subset of the visible properties present. Accordingly, when presented

with a visual scene consisting of a green triangle off to the right and a red square off to the left, if one has a visual sensation with mental qualities pertaining to a green triangle off to the right but no mental qualities pertaining to a red square off to the left, one will see the green triangle off to the right and fail to see the red square off to the left. And it could be that one has such sparse visual representations, but experience seems to present one with a detailed visual scene because one implicitly understands that if one moves one's eyes, one will see more details that one currently sees, as Noë claims.

Representations often represent things in terms of only a small subset of their properties. In fact, representations rarely, if ever, represent all of the properties of an object. Most drawings and paintings, e.g., are too coarse-grained to capture every visible wrinkle and pore of a person's face. And cartoons represent people and objects while leaving out much of the visible texture and 3-D shape of those objects.

Likewise, one's visual representations doubtless fail to represent all of the details of the objects they are representations of. And it could be that such visual representations leave out considerably more detail than we ordinarily think we see at a given moment. In this case, it could be that one fails to notice certain changes during change-blindness experiments because, though one sees the visual scene in virtue of forming a visual sensation with mental qualities pertaining to perceptible properties of stimuli, that visual sensation lacks the mental qualities that pertain to the perceptible properties and features

undergoing the change. If one's sensations fail to represent the changing features in a scene because those sensations lack the corresponding mental qualities, one will fail to notice the change. So the view that the visual representations involved in seeing are sparse in representational detail could account for change blindness.

But, according to Noë, we need not countenance visual representations at all to explain how we see the sparse details that we do in fact see at each moment. Rather, Noë claims, in addition to explaining why visual experience seems to us to present a great amount of visible detail, his enactive approach to visual perception also accounts for how we see those sparse details that we do see at a given moment.

Noë adopts James Gibson's (1966, 1979) view that seeing a stimulus does not involve visual representations or sensations that represent the features of that stimulus. Gibson's rejection of visual representations is motivated, not by surprising visual phenomena such as change blindness, but by the problem of explaining how we see stimuli as having objective, perceiver-independent perceptible properties, e.g., three-dimensional shapes, despite the impoverished, perceiver-relative nature of visual stimulation. When one sees an opaque cubical object, e.g., one sees only those sides, edges, and vertices facing one; only those sides, edges, and vertices project an image on the retina. As one moves around the cubical object, one sees other sides, edges, and vertices, but still only those projecting a retinal image at that moment. So one never sees all of the

sides, edges, and vertices of the cubical object at once; one sees at most three sides, nine edges, and seven vertices at a time.<sup>28</sup> And though one sees different sides, edges, and vertices at different moments, and the visual stimulation one receives is constantly changing as one moves, one sees the object as having an invariant, three-dimensional shape. So we must explain how one sees such invariant properties, such as three-dimensional shape, despite the constant changes in visual stimulation.

According to traditional views of perception, seeing such invariant properties of stimuli involves vision's forming static visual representations of stimuli pertaining to the retinal images they project, and then inferring (Fodor and Pylyshyn, 1981; Helmholtz, 1867/1962; Rock, 1997) or computing (Marr, 1982) the invariant properties from those static representations. According to such views, one determines that a stimulus is, e.g., cubical, by inferring or computing its shape from such factors as the changes in one's visual representations caused by the movements of one's eyes relative to that stimulus, or by the movements of the stimulus relative to one's eyes.

But Gibson claimed that seeing invariant properties does not require such computations or inferences involving visual representations (1966, p. 2). Rather, he claimed, one sees invariant properties of objects by directly picking them up from the light entering one's eyes, and one directly picks up such invariant

---

<sup>28</sup> Of course, one can see more sides, vertices, and edges of a translucent cube. I'll discuss only cases of seeing opaque objects here.

properties from the light by moving one's sensory receptors, e.g., one's eyes, thus changing the sensory stimulation caused by the light (1966, p. 4).

Noë unpacks Gibson's view that vision is sensitive to the relations between movements of one's eyes and changes in stimulation in terms of one's implicitly understanding those sensorimotor correlations. According to Noë, seeing the invariant properties of objects, such as their shapes, depends on one's implicit, practical understanding of sensorimotor relations (2004, pp. 77-79). On this view, one sees an object as cubical, e.g., in virtue of understanding how one's movements, e.g., those of one's eyes, are changing the stimulation the object is causing.

But one need not move at all to see a stimulus or its properties. One often sees a cube, e.g., as a cube without moving in relation to the object. Moving relative to a stimulus can of course change how one sees that stimulus by enabling one to see parts of it that one failed to see earlier. But one does in fact see the object as having an invariant shape even before one moves one's eyes. So Gibson and Noë must explain how one sees the properties of a stimulus before one moves one's eyes. And they must also explain why our ability to see a stimulus without moving does not show that their view is false.

According to Noë, one can see a stimulus as having a particular shape, e.g., without moving one's eyes relative to the stimulus as long as one exercises one's understanding of how one's movements could change the visual stimulation caused by the stimulus. On this view, one can see a stimulus as

cubical if one implicitly understands the ways in which moving one's eyes will change the sensory stimulation from its current state. One's implicit understanding of such sensorimotor correlations presumably depends on one's having seen a cube before, and on one's having moved one's eyes to visually explore cubes on such occasions. But once one understands how moving one's eyes relative to a cube transforms the visual stimulation caused by cubes, one can exercise such understanding even without moving one's eyes. When one sees the sides, edges, and vertices of a cube facing one, one sees the stimulus as cubical even without moving one's eyes because one implicitly understands how various movements of one's eyes relative to the stimulus will change the visual stimulation that the stimulus is causing (2004, p. 77).

Noë claims this view also explains one's perceiving properties other than shape. For example, according to this enactive view, seeing something off to the left is a function of one's grasping certain sensorimotor correlations, e.g., that moving one's eyes to the left will bring the stimulus into clearer view towards the center of one's field of view. Likewise, seeing something as far away is partly a function of understanding that if one moves forward, that object will in some way occupy an increasingly larger part of one's field of view, i.e., it will subtend an increasingly greater visual angle, so it will project an increasingly larger retinal image.

## 5. Dissociations of Visual Perception and Action

The view that perception is intricately linked to one's movements is of course not at all radical. We often suppose that it is because one visually perceives the location, shape, and size of an object that one is able to reach for and grasp that object. And folk psychology arguably individuates visual experiences in part by their causal connections to both motor inputs and motor outputs. Visual experiences of cubes, according to folk psychology, are states that are normally caused by cubes in good lighting conditions and that cause one to reach for and grasp cubical objects when one desires to grasp such objects. Folk psychology individuates such experiences in respect of other typical effects as well, such as their causing one to believe that a cube is present. But visual experiences are arguably also individuated partly in respect of their causing certain visually guided actions.

Noë, however, goes beyond this folk psychological claim that visual experiences often cause one's movements. According to Noë's enactive view, visual experience is constitutively linked to certain movements.

However, experiments in cognitive neuroscience suggest that separate visual processing streams underlie visual perception and visually guided action.<sup>29</sup> And Ned Block (2005) claims that these experiments pose a problem for Noë's enactive view of visual perception. Noë, according to Block, holds that there is a

---

<sup>29</sup> For a useful philosophical discussion of the view that visual perception and visually-guided action rest on separate processing streams, see Andy Clark (2001).

constitutive link between visual perception and visually guided action. So, if perception and visually guided action rest on separate processes, Block argues, Noë's view of perception fails.

Some of the most striking support for the claim that vision involves separate processing streams for visual perception and visually guided action comes from experiments on patients with the neurological disorders visual form agnosia and visual ataxia. Patients with visual form agnosia, which is caused by bilateral occipital lesions destroying the ventral prestriate cortex and disconnecting the inferior temporal lobes from visual input, are unable to report the orientations and shapes of visual stimuli. However, these patients can successfully perform visually guided actions, such as reaching and grasping stimuli presented in their visual fields, indicating that they do in some way see, or at least visually process, the orientations and shapes of those stimuli.

David Milner and Melvyn Goodale (1995) presented visual agnostic DF with a slot the orientation of which they varied between 0, 45, 90, and 135 degrees across trials. Milner and Goodale instructed DF to report the orientation of the slot, and then to insert either her hand or a note card into it. They found that though DF could not report the orientation of the slot, she could easily orient her hand or a note card to insert it into it. They take this to show that though DF could not visually perceive the orientation of the slot, she visually processed the information about the slot in a way that enabled her to perform the visually guided action. And Milner and Goodale conclude that this shows that vision

involves two separate processing streams, the ventral stream, which underlies visual perception and is damaged in DF, and the dorsal stream, which underlies visually guided action and remains intact in DF.

But, one might argue, this experiment does not show that visual perception and visually guided action rest on separate processing. Rather, it could be that DF fails to report the orientation of the slot, not because she fails to perceive it, but because she fails to perceive it consciously. One reports only what one perceives consciously. But nonverbal, overt behavior also reflects what one perceives. So it could be that DF successfully orients her hand to fit it into the slot because she visually perceives its orientation, but she fails to report the orientation of the slot because she does not perceive it consciously.

This view rests on the assumption that conscious perception involves a process by which one perceives stimuli and another process by which one is conscious of perceiving those stimuli, and that the process in virtue of which one is conscious of perceiving, but not that in virtue of which one perceives, is damaged in DF. If so, conscious perception requires both the dorsal processing stream, which remains intact in DF, and the ventral processing stream, which is damaged in DF.

If this view is correct, one could not consciously perceive a stimulus without having dorsal processing, even if one's ventral processing stream is intact. So one could not consciously perceive the orientation of a stimulus without the processes underlying one's ability to act on that orientation, unless of

course one has some motor deficiency preventing one from performing such actions.

But experiments on patients with optic ataxia, which results from damage to the dorsal processing stream, suggest that the dorsal processing underlying visuomotor actions, such as those exhibited by DF, are not identical with unconscious perception. Rather, these experiments confirm Milner and Goodale's theory that visual perception and visually guided action rest on separate processing.

Patients with optic ataxia exhibit dissociative behavior opposite to that exhibited by patients with visual form agnosia. Optic ataxics successfully report the orientations of stimuli, such as slots, but they fail to accurately orient their hands to insert them into those slots (Perenin and Vighetto, 1988). Optic ataxics thus exhibit intact perception of orientation, but disrupted visuomotor skills. If visually guided action, such as hand orienting, in the absence of conscious perception is best explained in terms of one's unconsciously perceiving orientation, then one who consciously perceives the orientation of a slot could arguably both report the orientation of the slot and orient one's hand to fit it into the slot. However, optic ataxics consciously perceive the orientations of slots, as reflected by their accurate reports of the orientations of the slots, but they fail to orient their hands to insert them into the slots.<sup>30</sup> This shows that the processing

---

<sup>30</sup> Optic ataxics can perform motor actions such as reaching and grasping, as long as it is not visually guided. So their inability to perform visually guided movements does not result from a motor deficiency.

underlying visually guided action is dissociable from the processing involved in conscious visual perception. So the processing involved in visually guided action is not required for conscious perception. This in turn suggests that DF's ability to accurately orient her hand to insert it into a slot does not depend on intact, unconscious perception of orientation.

Taken together, the occurrence of visual form agnosia and optic ataxia show that visual perception and visuomotor skills are doubly dissociable, suggesting that the visual processing underlying visual perception and the visual processing underlying visually guided action are distinct.

Support for a dissociation between visual processing underlying visual perception, on the one hand, and visuomotor processing, on the other, is not limited to cases involving subjects with neurological disorders. Rather, there is also evidence for such dissociations in normal subjects.

For example, Goodale and Kelly Murphy (1997) showed that subjects perceive the sizes of visual stimuli presented foveally more accurately than they perceive the sizes of visual stimuli presented peripherally, but visual processing for visually guided action towards peripherally presented stimuli is not less accurate than visual processing for visually guided action towards foveally presented stimuli.

Goodale and Murphy presented subjects with both a perceptual task and a visuomotor task. In the perceptual task, subjects were first presented with 5 blocks of different widths. Subjects were trained to rank the widths of these

blocks on a scale from 1 to 5. In each trial of the test phase, subjects were presented with a block at a position ranging from 5 to 70 degrees from the line of sight, and they were to report the width of the block using the scale they learned in the training session.

In visuomotor trials, subjects were presented with a block positioned between 5 and 70 degrees from the line of sight, and they were asked to reach for and grasp the block. As subjects reached for the blocks, Goodale and Murphy measured the maximum aperture of subjects' grips, measured as the distance between the subject's index finger and thumb.

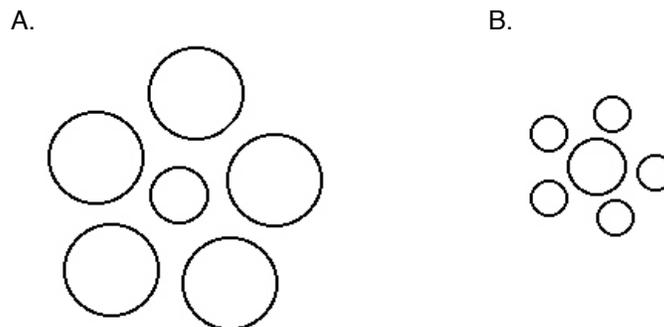
Goodale and Murphy found that subjects in the perceptual task accurately reported the widths of blocks presented towards the fovea and underestimated the sizes of the blocks presented towards the periphery. However, in the visuomotor task the subjects' grips were accurately scaled to the widths of the blocks no matter where the blocks were presented.

These results suggest that one's ability to perceive the sizes of stimuli is significantly more accurate in foveal vision than in peripheral vision, but one's ability to accurately adjust one's grip aperture to grasp objects does not depend on where in one's field of view the stimulus is presented. This in turn supports the view that visual perception and visually guided action depend on separate processes.

Other experiments show that certain visual illusions affect visual perception but not visually guided action. For example, Angela Haffenden, Karen

Schiff, and Goodale (2001) showed that the Ebbinghaus illusion affects subjects' perception of the sizes of stimuli but it does not affect their grasp scaling when subjects reach for those objects. The Ebbinghaus illusion is a visual illusion whereby two circles of equal diameters look different in size when one is surrounded by an annulus consisting of small circles and the other is surrounded by an annulus consisting of larger circles. A circle surrounded by an annulus of small circles looks bigger than a circle of the same diameter surrounded by an annulus of larger circles.

Figure 5: The Ebbinghaus Illusion



The diameter of the central circle in A is equal to the diameter of the central circle in B, even though the central circle in B looks larger than the central circle in A.

Haffenden et al. presented subjects with two circles of equal diameters, one surrounded by an annulus of small circles and the other surrounded by an annulus of larger circles. They asked the subjects to manually estimate the size of the target circles using the distance between their thumbs and index fingers to estimate the diameters of those central circles. Haffenden et al. then asked

subjects to reach for the central circles to grasp them, and measured subjects' grip apertures as they reached for the central circles. They found that subjects' manual estimates of the sizes of those circles reflected the Ebbinghaus illusion; subjects used a wider grip aperture to estimate the size of the circle surrounded by an annulus of small circles than they used to estimate the size of the circle surrounded by an annulus of larger circles. However, Haffenden et al. also found that subjects' grip apertures when reaching for the circles do not reflect the Ebbinghaus illusion; subjects' grip apertures were the same when reaching for the circle surrounded by an annulus of small circles and when reaching for the circle surrounded by an annulus of larger circles. This result further supports the view that visual perception and visuomotor action depend on distinct processes.

According to Block, such dissociations between visual perception and visually-guided action show that Noë's enactive approach to perception fails (2005, pp. 268-269). According to Noë, perception depends on one's understanding correlations between visual stimulation, on the one hand, and one's movements, on the other. But, Block argues, perception does not rest on one's understanding visuomotor correlations, since the visual processing underlying one's visually guided movements is distinct from the visual processing underlying perception. If Noë's view were correct, Block assumes, dissociations between visual perception and visually guided movement would not arise.

But the dissociations between visual perception and visuomotor processing do not show that Noë's enactive approach to perception is false. Noë

holds that visual perception rests on one's implicitly understanding how one's movements affect visual stimulation. Accordingly, to see the shape, size, or location of an object, e.g., one must understand how moving one's eyes in such-and-such a way will change the stimulation caused by the stimulus, or how movements of the stimulus relative to one's eyes will change the stimulation that stimulus produces. But Noë does not claim that visual perception requires an understanding of how one can act on stimuli one sees.

The bodily movements involved in such visually guided actions as reaching and grasping are outputs of visual processing; they are bodily movements caused by processing of visual information, such as orientation, shape, and location. But the movements Noë invokes in his account of visual perception are inputs to visual processing. Noë claims that visual perception constitutively depends on the inputs provided by, e.g., eye movements. He does not claim that visual perception constitutively depends on one's ability to act on visual stimuli.<sup>31</sup>

Of course, the sensorimotor understanding Noë invokes does include one's understanding of how visual stimulation varies with movements of one's limbs, head, and sometimes one's entire body; it is not limited to movements of one's eyes. For example, when one explores a cubical object by turning it with one's hand to view the object's previously hidden features, according to Noë's

---

<sup>31</sup> Further, optic ataxics successfully direct their eyes towards objects they cannot reach for or grasp (Riddoch, 1935; Ratcliff and Davies-Jones, 1972). This suggests that it could be that some visually guided movements are in fact involved in visual perception, even when other visuomotor abilities are absent.

view, one sees it as cubical in virtue of understanding how those hand movements relate to changes in the stimulation caused by the object. But in those cases, kinesthetic information about one's hand movements serve as inputs to visual processing, they are not outputs of visual processing. Indeed, one could visually explore an object without moving one's body at all. If the object itself is rotating and one is aware of its movements relative to one's eyes, one could determine how the object's movements are changing the visual stimulation.

The experiments Block cites against Noë's enactive view show only that visual perception and visually guided action rest on separate processing. Those experiments do not show that one perceives, e.g., an object's shape in the absence of inputs about one's movements relative to the object. So Block's argument against Noë's enactive approach to perception fails.

Of course, even if Noë did argue for a constitutive link between seeing and visually guided actions, the data showing that separate visual-processing streams underlie visual perception and visuomotor activity would not argue against his view. Although certain visually guided actions do not depend on visual perception, there are of course actions that are guided by what one visually perceives. Visual perception could be constitutively linked to those perceptually guided actions, even if it is not constitutively linked to other visually guided actions, i.e., those resulting from dorsal processing.

## 6. The Perspectival Character of Seeing

Although Noë's view withstands Block's objection, as stated it does not fully capture the qualitative character of seeing. When one sees a cube without moving in relation to it, one often sees it as cubical, even though one sees it from only one perspective. And when one sees a cube from a different perspective, one can also see it as a cube, again even if one does not move in relation to the cube. But seeing a cube from one perspective is qualitatively different from seeing a cube from another perspective. And that qualitative difference is arguably independent of one's seeing the stimulus as cubical, since one sees the cube as cubical from these different perspectives. Noë must account for such qualitative differences.

Noë must also explain the qualitative similarities that hold between seeing different shapes from different perspectives. For example, seeing an object with a circular surface, e.g., a penny, tilted at an angle away from one's eyes is qualitatively similar to seeing an object with an elliptical surface perpendicular to one's line of sight. Nevertheless, one usually sees such tilted circular objects as circular, not as elliptical. And one usually sees elliptical objects as elliptical, not circular.

So the qualitative character of seeing shape is not exhausted by the invariant shape one sees a stimulus as having. Noë must explain how his enactive approach accounts for this further aspect of visual experience.

Noë claims that when one sees an object one sees not only its invariant perceptible properties, e.g., its three-dimensional shape, but also what he calls its perspectival properties. And he holds that seeing those perspectival properties accounts for the perspectival aspect of seeing. Perspectival properties are properties a stimulus has in virtue of its bearing certain spatial relations to a perceiver's body (2004, p. 83). For example, the perspectival size of a stimulus is a property corresponding to the size of the region of a plane perpendicular to one's line of sight that one would have to fill to occlude the stimulus and nothing else. This property is distinct from the perceiver-independent, invariant size of the stimulus, since two stimuli of different sizes could have the same perspectival size when they are positioned at different distances from the perceiver. Likewise, two stimuli of the same size could have different perspectival sizes when they are positioned at different distances from the perceiver. Though such perspectival sizes are properties stimuli have in relation to the perceiver, they are not subjective, mental properties of perceptual states. Rather, perspectival sizes are relations between stimuli and one's perceptual system.

Stimuli have perspectival shapes as well. A circular object, such as a penny, has one perspectival shape when its surface is perpendicular to one's line of sight and it has a different perspectival shape when its surface is tilted 45 degrees from one's line of sight. These perspectival shapes correspond to the shape of the retinal image the penny projects; a penny with its surface

perpendicular to one's retina projects a circular retinal image, and a penny tilted 45 degrees away from one's retina projects an elliptical retinal image. Again, these properties depend on the spatial relations the stimulus bears to the perceiver, but they are nonmental, perceptible properties of those stimuli.

One might argue that perspectival properties are not legitimate physical properties, but perceiver-dependent properties countenanced only to preserve a particular theory of perception.<sup>32</sup> But there are reasons independent of Noë's theory, as well as the very different theory of perception I argued for in chapter 1, to countenance perspectival properties.<sup>33</sup> In addition to producing different perceptual experiences, a penny tilted 45 degrees from a surface produces different nonperceptual effects from those that a penny positioned parallel to that surface produces. If the surface is reflective, the tilted penny causes an elliptical reflection, but the penny positioned parallel to the reflective surface produces a circular reflection. Likewise, a tilted penny casts an elliptical shadow on a surface, whereas a penny positioned parallel to that surface produces a circular shadow, assuming that the light source is directly behind the penny. So the tilted penny and the penny positioned parallel to the surface differ in respect of some properties in virtue of which those pennies cause differently shaped reflections and shadows. Neither reflections nor shadows are mental phenomena, and they

---

<sup>32</sup> Gary Hatfield raised this objection to such properties during a 2/15/06 talk to the CUNY Graduate Center Philosophy Colloquium.

<sup>33</sup> Sydney Shoemaker (1996) and Michael Tye (1996) also offer distinct views of perception that countenance perspectival properties.

exist independently of one's perceiving them. So the properties of the pennies that produce differently shaped reflections and shadows are posited to explain, not just the various ways pennies look to us when seen from different perspectives, but a number of effects such pennies produce independently of perception. Those properties are perspectival properties, properties one perceives in virtue of seeing stimuli from a particular perspective.

According to Noë, such perspectival properties account for the qualitative character of seeing stimuli from different perspectives. When one sees a penny straight on, one sees a particular perspectival shape; when one sees the penny at an angle, one sees another perspectival shape. Further, the penny seen at an angle from one's line of sight and an elliptical object seen straight on look similar to each other, according to this view, because they share a perspectival shape.

Perceiving such perspectival properties is integral to Noë's enactive approach to perception. Although, prior to his discussion of perspectival properties, Noë characterizes seeing in terms of one's understanding the correlations between one's movements and changes in visual stimulation, he later modifies that view to accommodate the perspectival character of seeing. According to the modified version of Noë's view, we see invariant properties, such as 3D shapes, in virtue of exercising implicit, practical understanding of the ways the perspectival properties of a stimulus vary with the movements of one's eyes relative to the stimulus (2004, p. 84). One exercises such sensorimotor understanding when one visually explores a stimulus, i.e., when one moves

one's eyes relative to the stimulus and determines how the perspectival properties one sees change as a result of those movements. And one also sees invariant properties without moving one's eyes, i.e., when one sees the stimulus and implicitly understands how moving one's eyes would change the perspectival properties of the stimulus. So to see a stimulus as, e.g., cubical, one must see the perspectival shape of the stimulus.

If Noë's enactive approach to seeing invariant properties depends on one's seeing perspectival properties, then Noë must account for how one's sees the perspectival properties of stimuli.

Noë's account of how we see invariant properties of stimuli does not apply to our seeing perspectival properties. If it did, one would see a perspectival shape, e.g., by seeing some property P that is distinct from both the perspectival shape and the invariant shape of the stimulus, and implicitly understanding how moving one's eyes would change property P to some other property that is also distinct from any perspectival shape or invariant shape of the stimulus. Such an account would thus be committed to one's implicitly understanding how movements of one's eyes change nonperspectival, noninvariant properties of the stimulus. But such a view must then explain both the nature of those properties and how we see those properties. Any attempt to do so in terms of further sensorimotor understanding leads to a regress of perceptible properties of stimuli and sensorimotor correlations.

Further, the enactive approach to explaining how one sees stimuli as having viewpoint-independent, invariant properties is supported by one's ability to see stimuli as having such invariant properties despite one's failing to see all of the parts of a three-dimensional stimulus at once. One must explore a cube, e.g., or at least understand how to move one's eyes in order to explore it, to see the cube as a cube because one sees only those sides, angles, and vertices of the cube facing one. But seeing a perspectival shape does not require such active exploration; one sees all of the perspectival shape at once, from a single perspective. So seeing a perspectival property arguably does not require any understanding of the ways movements of one's eyes change any other perceptible properties of a stimulus.

Nevertheless, Noë does attempt to explain how we see perspectival properties in terms of our implicit understanding of sensorimotor correlations. According to Noë, one sees a particular perspectival property, e.g., a perspectival shape, "... only insofar as, in encountering it, one is able to draw on one's appreciation of the sensorimotor patterns mediating (or that might be mediating) your relation to it. How you appreciate it as being is constituted by the sensorimotor knowledge you bring to bear in your encounter with it" (2004, p. 90). According to this view, seeing the perspectival properties of an object, like seeing its invariant properties, depends on one's implicit understanding of sensorimotor correlations.

But it is not clear how one could bring any sensorimotor understanding to bear on a perspectival property if one did not already see that perspectival property. And it is not clear how one could see an invariant property in virtue of understanding how the perspectival properties one sees would change as a result of one's movements if seeing those perspectival properties itself depends on one's understanding how they would change as a result of one's movements.

Further, the sensorimotor understanding one brings to bear on the perspectival shape one sees when one sees, e.g., an elliptical object perpendicular to one's line of sight is different from the sensorimotor understanding one brings to bear on the perspectival shape one sees when one sees a circular object tilted from one's line of sight. According to Noë, it is because one brings different sensorimotor understanding to bear in these situations that one sees the first stimulus as elliptical and the second as circular. But, according to Noë, these stimuli share the same perspectival shape. So seeing a perspectival shape depends on something other than the sensorimotor understanding one brings to bear on that perspectival shape. So it is unclear how Noë's view could explain how one sees the perspectival properties of objects.

But we can explain how we see perspectival properties in terms of the view that we have visual sensations that represent them. According to this view, one sees the perspectival shape of a penny tilted away from one's eyes in virtue of having a visual sensation with a mental quality pertaining to that perspectival

shape. And that mental quality is similar to the mental quality of the visual sensation one has when one sees an elliptical object straight on. This is the view I proposed in chapter 1. There I argue, against Peacocke (1983, 2001), that such mental qualities represent objective, perceptible properties of stimuli, i.e., perspectival properties, and they account for the qualitative character of seeing such properties. But if we see perspectival properties in virtue of having visual sensations that represent them, then Noë's enactive view of how we see invariant properties rests on our having visual representations.

In any case, there is further support for the claim that Noë's enactive approach to visual perception depends on visual representations of the properties of stimuli. According to the enactive account, one sees a stimulus as, e.g., cubical in virtue of implicitly understanding how movements, such as movements of one's eyes relative to the stimulus, will change the perspectival shapes one sees. And this implicit understanding of such sensorimotor correlations presumably rests on one's having visually explored cubes before.<sup>34</sup> When one visually explores a cube, one moves one's eyes around it, thus changing the perspectival shapes one sees, while also changing which sides, edges, and vertices one sees. According to Noë, seeing the stimulus as a cube requires that one is aware of the correlations between the movements one has made and the changes in perspectival shapes one has seen. So at one moment, one must be aware of how one has just moved one's eyes in relation to the

---

<sup>34</sup> I am assuming here that one's understanding of sensorimotor correlations is not innate.

stimulus and one must be aware that that particular movement resulted in a change in the perspectival shape from the perspectival shape one saw a moment earlier to the perspectival shape one is currently seeing. To see such a change in the perspectival shape, and to be aware of how that change relates to one's movements, one must remember what perspectival shape one saw a moment earlier. But remembering the perspectival shape one previously saw requires a persisting representation of that perspectival shape, since one is no longer looking at that perspectival shape. So visually exploring a stimulus to perceive its invariant shape requires representations of perspectival shapes.

Nevertheless, Noë might argue that this does not show that all seeing involves visual representations. As discussed earlier, one often sees a stimulus without moving one's eyes. In this case, one sees perspectival properties of the stimulus, but one also sees the stimulus as having nonperspectival, invariant properties. Seeing the stimulus as having those invariant properties, according to Noë, is a function of one's understanding how moving one's eyes would change the perspectival properties of the stimulus if one moved one's eyes, but one need not move one's eyes in order to understand those sensorimotor correlations. In such cases, Noë could argue, seeing the stimulus as having those invariant properties does not require any visual representations at that time. Rather, one simply sees the perspectival properties and implicitly understands how moving one's eyes would change those perspectival properties if one were to move one's eyes.

But, if visual representations of perspectival properties are required for one to develop an understanding of the sensorimotor correlations involved in one's seeing stimuli as having invariant properties, and if they are required for one to determine what invariant properties of a stimulus one is in fact visually exploring, then it is not clear why Noë would deny that one sees perspectival properties of stimuli in virtue of having visual representations of those perspectival properties.

In fact, if visual representations are required for visually exploring objects, but they are not involved in our seeing perspectival properties, it is not clear how vision could generate the visual representations when one visually explores an object. If seeing involves visual representations of perspectival properties only when one is visually exploring an object, then vision must somehow predict that one will move one's eyes to explore an object in order to determine when to generate a visual representation. Without generating visual representations of the perspectival properties of a stimulus before one moves one's eyes, vision could not subsequently determine how the perspectival properties of the stimulus changed as a result of that eye movement.

Perhaps Noë could argue that vision uses feedback signals from motor commands<sup>35</sup> to determine when one will move one's eyes, and it then generates a representation of the perspectival properties one is currently seeing before one's eyes move.

---

<sup>35</sup> Patrick Haggard (2005), e.g., argues that there is such feedback from motor commands.

But this complex procedure does not help explain how one determines the invariant shape, e.g., of an object that is rotating on its own. In that case, there are arguably no motor commands that could trigger vision to generate a visual representation. So determining the invariant shape of an object on the basis of correlations between the object's movements relative to one's eyes and the changes in the perspectival properties that the object presents requires that vision already has generated visual representations of the object's perspectival properties. The best explanation is that visual perception of perspectival properties involves visual representations of those properties.

Further, visual representations best explain how we see changes in visual scenes when we see those changes. Even during change-blindness experiments, subjects often successfully notice what feature of the scene changed. For example, half of the subjects in Grimes's experiments noticed that the heads of two cowboys in a picture switched places, even though that switch occurred while the subjects saccaded. But to see that a particular feature has changed, one must remember the features of the scene before the change. Noticing that change requires that those subjects remembered which head was located where before their saccades. And such memory requires a representation of at least one of the heads at its previous location. It is unclear how one could notice such changes if one did not have a visual representation of the changed feature in the first place, just as it is unclear how one could learn

sensorimotor correlations or apply them while visually exploring a stimulus if one did not have such representations.

Since noticing changes in a visual scene and perceiving the invariant properties of objects by visually exploring them are best explained in terms of the view that visual perception involves representations, we can best explain how one sees perspectival properties in terms of one's having visual representations of those properties.

This suggests Noë's view that visual perception involves no visual representations of the features of visual scenes is wrong. However, it does not show that he is wrong that we see very little of a visual scene at each moment. Perhaps Noë is right that change blindness shows we see very little at each moment. If so, since seeing requires visual representations, such as sensations, perhaps change blindness shows that such representations represent only a small subset of the visible features of a scene at each moment. So perhaps change blindness shows that visual representations, such as visual sensations, are sparse in representational detail, not that visual perception involves no representations, as Noë argues.

There is independent support for the view that the representations involved in visual perception are in fact sparse in representational detail. Visual acuity is much better at the fovea than at the periphery. This is because there are many more retinal receptors at the fovea than there are at the periphery (Grimes, 1996, p. 90). So, without moving one's eyes, one will see very little at

once; one will see the details of stimuli presented foveally, but one will see much less detail of stimuli presented peripherally. However, one's eyes saccade a number of times each second, enabling one to shift one's fixation from foveal objects to peripheral objects. So visual perception is rarely if ever significantly limited by the poor visual acuity at the periphery. In fact, we are rarely even conscious of our saccades, so we are rarely conscious of how little we see without saccading.

But the sparseness of fixation-dependent visual representations does not by itself show that visual sensations are themselves sparse in detail. It could be that vision constructs visual sensations out of consecutive, momentary, fixation-dependent visual representations of the parts of the visual scene one sees during each fixation. If so, it could be that those fixation-dependent representations are sparse, but sensations are highly detailed.

Nevertheless, if visual sensations are in fact constructed out of sparse, fixation-dependent subpersonal representations, perhaps change blindness is in fact best explained in terms of the failure of one's sensations to represent all of the detail of a visual scene after all. If the fixation-dependent subpersonal state that represents a changing feature fails to be integrated with the other subpersonal representations comprising one's visual sensation, then one's visual sensation will fail to represent that changing feature. And, if one's sensation fails to represent that feature, one will fail to see the change.

For example, consider Grimes's experiment in which subjects fail to report a significant change in the size of a single prominent building in a picture of a city skyline. It could be that subjects fail to see that change because when their visual systems constructed their visual sensations, they failed to integrate a subpersonal, fixation-dependent representation of the changing building either before or after the building changed size. If so, the subjects' visual sensations of the picture failed to represent that changing building, so the subjects failed to see the change.

So it could be that change blindness results from one's sensation's failure to represent a changing or changed feature, even if one's sensations are considerably more detailed than the fixation-dependent subpersonal representations that comprise them. Perhaps Noë is right that change blindness results from one's failure to see a changing feature, even if he is wrong that change blindness shows that we do not see visual scenes in virtue of having visual representations of those scenes.<sup>36</sup>

---

<sup>36</sup> However, as I discuss in the next chapter, there are several other accounts of change blindness that do not advert to the representational sparseness of sensations. There I argue that at least some cases of change blindness occur even when one does see the changing features, so even when one's sensations represent those features.

## 7. Homomorphism Theory and Sparse Sensations

Homomorphism theory, the view of the qualitative character of sensing I argued for in the previous chapter, is compatible with this account of change blindness.

According to homomorphism theory, the visual representations, or sensations, in virtue of which we see visual stimuli and scenes have mental properties, i.e., mental qualities, that pertain to the perceptible properties of those stimuli and visual scenes. Accordingly, visual sensations have mental qualities that represent the spatial properties of stimuli and to the spatial layouts of visual scenes.

The mental qualities of visual sensations pertaining to the spatial properties of visual stimuli and scenes represent those spatial properties in virtue of homomorphisms between families of those mental qualities and families of perceptible spatial properties. A visual sensation of a square, e.g., has a mental quality, square\*, that resembles and differs from other such mental qualities in ways parallel to the ways perceptible squares resemble and differ from other perceptible shapes. Just as perceptible squares are more similar to perceptible rectangles than to perceptible triangles, the mental quality square\* is more similar to the mental quality rectangular\* than to the mental quality triangular\*. Likewise, homomorphisms hold between families of other mental qualities and families of other perceptible spatial properties, e.g., visible sizes, visible locations within one's field of view, and visible orientations.

It is a theoretical claim that we have visual sensations with such mental qualities, and that those mental qualities represent perceptible spatial properties by way of homomorphisms. According to homomorphism theory, sensations and mental qualities are theoretical posits posited to explain how we sense the perceptible properties of stimuli and scenes. We can see the difference between a square and a triangle. And we can best explain how we see this difference in terms of the view that we have visual mental states that differ in respect of some properties that pertain to those different perceptible shapes. Likewise, we see relative similarities between different shapes; squares look more similar to rectangles than triangles. And we can best explain how we see such relative similarities in terms of the view that our visual states have properties that bear similar resemblance relations. So homomorphism theory accounts for the ways we perceptually discriminate various perceptible properties, and for how we see the perceptible properties we do in fact see.

But homomorphism theory is not committed to the view that our visual representations, or sensations, are highly detailed. It is committed only to the view that one's current visual sensations have mental qualities that pertain to each perceptible property one sees, i.e., visually perceives, at that time. But it could be that while one sees a scene consisting of a city skyline with five rectangular buildings, e.g., one's visual sensation does not have mental qualities pertaining to all of those rectangles. Perhaps one's visual sensation has mental qualities, i.e., shapes\*, pertaining to only four of those rectangles. In that case,

one will see only four buildings.<sup>37</sup> Of course, one's behavior will reflect one's failure to see the fifth building. For example, one will not comment on that building, nor will one correctly count the number of buildings in the picture if asked to. Further, if the picture of the building that one's sensation fails to represent changes while one is looking at the picture, one will not see those changes. And when one is asked if one saw the changes, one will report that one did not. So homomorphism theory is compatible with the view that change blindness results from our having sparse visual representations, or sensations.

---

<sup>37</sup> This does not of course rule out that one's visuomotor processing stream encodes the information about the building that one's visual perceptual system fails to represent, since sensations are states involved in visual perception, not visuomotor processing.

## Chapter 3: Change Blindness, Part 2

### 1. Introduction

In the previous chapter, I argued against Noë's claim that change blindness poses a problem for the traditional view of visual perception, according to which visual perception involves visual representations, or sensations, of the spatial layouts of visual scenes. I argued that Noë's enactive view, the view he proposes as an alternative to the traditional view, itself requires visual representations. In addition, I argued that the view that visual perception involves sparse visual representations that fail to represent the changing features of visual scenes could account for change blindness, and that the homomorphism theory of sensing I argued for in Chapter 1 is compatible with that account.

In this chapter, I will examine alternatives to the view that change blindness is a failure to see changing features that results from a failure of visual representations, or sensations, to represent the changing features of a visual scene. I will focus primarily on Fred Dretske's (2004) view that change blindness is a failure to see that a visual scene is changing, but it is not a failure to see the changing features themselves. Dretske's account of change blindness rests on his claim that subjects' reports that they do not see any changes, and their failure to report changes, do not show that those subjects fail to see the changing features. If so, change blindness does not show that subjects' visual

representations of visual scenes represent only a small subset of the features of those scenes.

I'll argue that change blindness could occur even when subjects do see that a visual scene has changed, since they could see that change without being conscious that they see it. I'll then examine psychological and neuroscientific experiments that arguably support the view that change blindness is due to one's failure to be conscious that one is seeing a change, not to one's failure to see that change. Finally, I'll argue that the homomorphism theory of sensing is compatible with this account of change blindness, as well as with a number of alternative accounts.

## 2. Change Blindness Despite Detailed Visual Sensations

The view that change blindness results from one's having sparse visual representations rests on the assumption that one does not see the changing features of the visual scene, or at least that one fails to see the changing feature either before or after the change. But, one might argue, change blindness does not show that if visual perception involves visual representations, such as sensations, of the spatial layouts of visual scenes, those representations are sparse in representational detail. The main motivation for holding that our visual representations are sparse in representational detail is that subjects fail to notice significant changes during change-blindness experiments. But it could be that subjects see visual scenes in virtue of having very detailed visual representations

that do represent the changing features, but not all of the details of those representations are subsequently encoded in short-term, working memory (Rensink, O'Regan, and Clark, 1997; Rensink, 2000). If so, it could be that one fails to notice a change in a feature that one did in fact see, i.e., a feature that one's visual sensation represented, because vision failed to store a representation of the feature that subsequently changed. Details that fail to enter into working memory fail to affect further visual processing, such as change detection. So subjects' failure to notice changes in visual scenes could be due to such failures of memory, not failures to see changing features. So the view that visual perception involves the formation of detailed visual representations, but that visual working memory stores only a limited subset of those represented details, could also account for change blindness.

Alternatively, it could be that we see visual scenes in virtue of having visual sensations that represent a great amount of detail, and much of that representational detail is stored in working memory, but not all of the stored representational detail is then compared with features of newer visual representations (Mitroff, Simons, and Levin, 2004). On this view, one fails to notice changes in a visual scene when vision fails to compare a stored representation of the changed feature with a new representation of that changed feature. Accordingly, one fails to notice that the visual scene has changed because whatever mechanism compares visual representations from moment to moment fails to compare representations of the relevant, changing features.

### 3. Verbal Reports and Change Blindness: Dretske

Dretske (2004) offers yet another account of change blindness that is also compatible with the view that we see visual scenes in virtue of having highly detailed visual representations of those scenes.<sup>38</sup> According to Dretske, subjects in change-blindness experiments do in fact see the changing features, but they do not see those changing features in a way that makes them aware that the features have changed. In this respect, Dretske's view, like the view that change blindness results from a failure to compare the changing details of visual representations pertaining to the changing features of a scene, and like the view that change blindness results from vision's failure to store representations of the features that change, provides an alternative to the view that change blindness shows that our visual representations of a scene represent less detail than we ordinarily think they do. The view that change blindness shows our visual representations are sparse in representational detail rests on the assumption that subjects in change-blindness experiments fail to see the changing features of the scene. If Dretske is right that change blindness does not result from a failure to see the changing details, it does not show that the visual representations in virtue of which we see visual scenes are sparse.

---

<sup>38</sup> Though Dretske's view appears to be an alternative to the views I have already discussed, it could be that it is in fact a version of the view that change blindness results from some sparse visual representation, the view that change blindness results from a memory failure, or the view that change blindness results from vision's failure to compare visual representations of the changing features. I will discuss these versions of Dretske's view towards the end of this section.

According to Dretske, the view that change blindness shows we often fail to see changing features of visual scenes rests on a faulty inference from subjects' reports. Subjects in change blindness experiments fail to report the features changing in a scene, e.g., they fail to report the change in size of a prominent building in a picture of a city's skyline. However, Dretske claims, this does not itself show that those subjects fail to see the changing features, since one's reports do not always reflect everything that one sees. Rather, Dretske argues, one's reports reflect what facts one sees, but they do not always reflect what objects or features one sees.

Dretske's account of change blindness rests on a distinction he draws between two ways of being aware of stimuli, one of which he calls fact awareness and the other of which he calls object awareness. The distinction between fact awareness and object awareness can be illustrated by many commonplace examples. Suppose that Jones sees her neighbor. And suppose that Jones's neighbor is a spy. However, suppose also that Jones does not believe that her neighbor is a spy. If Jones is asked what she sees, she will not say that she sees a spy. And if Jones is asked whether she sees a spy, she will say that she does not see a spy. Of course, though she does not believe she sees a spy, Jones does in fact see a spy, since she sees her neighbor who is in fact a spy. According to Dretske, although Jones sees a spy, and is therefore aware of a spy, she is not aware of the fact that the person she sees is a spy. In

Dretske's terminology, Jones is object aware of a spy, but she is not fact aware of a spy.

Dretske claims that because of this distinction between object awareness and fact awareness, we must be careful when drawing inferences about what one sees on the basis of what one verbally reports. Though one's verbal reports reflect one's fact awareness, he claims, such reports do not always reflect one's object awareness.

One can perceive—consciously perceive—spies and flying saucers (teapots, bicycles, etc.) while sincerely denying awareness of any such thing. Behavioral measures of consciousness that tie a person's perception ... of x in location L too closely to the person's ability to report his awareness of x in L tend to confuse conscious perception of objects ... with conscious perception of facts—either the fact that there is an x in L or the fact that one is aware of an x in L ... Although you can't see (the fact) that there are spies in the neighborhood without believing that there are spies in the neighborhood, you can certainly see spies in the neighborhood while believing that there are none (and, therefore, that you are aware of none). (2004, pp. 7-8; emphasis in original)

Subjects in change-blindness experiments often fail to report the changing features of scenes, they often fail to report that they see any changes at all, and they often deny that they see any changes. However, according to Dretske, just as Jones's failure to report seeing a spy and her failure to report the presence of a spy do not show that Jones fails to see a spy, subjects' failures to report changes in visual scenes during change-blindness experiments, and their denials that they see those changes, do not show that they fail to see those changing features. Those reports, according to Dretske, show only that the subjects are not aware that those features are changing. In Dretske's terminology, one's failure to report changes shows that one is not fact aware that a feature of the scene has changed, but it does not show that one fails to see the changing features, since one could be object aware of the changing features while being unable to report the changes. Since being object aware of a stimulus is insufficient for reporting that stimulus, one's failure to report what features are changing in a scene does not show that one failed to see those changing features. So, Dretske claims, change blindness experiments do not show that subjects fail to see the features that distinguish the visual scenes they are presented with.

Dretske does not explain how we determine whether one is object aware of something. Rather, he assumes that one is object aware of something if that thing is positioned in front of one's open, functioning eyes, unless one sincerely

reports seeing nothing at all (2004, pp. 8-9).<sup>39</sup> If so, subjects arguably have visual representations in virtue of which they are object aware of the changing features of a visual scene, but they fail to report those changes because the representations involved in object awareness are insufficient for reporting the things they make one aware of. In that case, one's failure to report changes in a change-blindness experiment does not show that one's visual representations fail to represent the changing features of the scenes. So change blindness does not show that the visual representations in virtue of which one sees visual scenes are sparse in representational detail.

But, one might argue, one's failure to report the changes in visual scenes shows that the visual representations involved in one's fact awareness of those changing features are sparse, i.e., that those visual representations do fail to represent the features that change. If at one moment a subject is fact aware that a prominent building in a picture is taller than all of the other buildings in the picture, and one is subsequently fact aware that that building is shorter than some of the other buildings, then the subject will presumably also be fact aware that the building has changed size between those two scenes. So, one might argue, when subjects fail to notice such changes, they must have failed to be fact aware of the size of the building in at least one of the scenes.

---

<sup>39</sup> This is compatible with the view that one is object aware of something if and only if one is also fact aware of it. On Dretske's view, Jones is object aware of a spy when she sees her neighbor, even though she is not aware of the fact that her neighbor is a spy. However, Jones is fact aware of her neighbor; she is aware of the fact that he is her neighbor, among other things.

However, it could be that at one moment one is fact aware that, e.g., a prominent building in a picture is taller than the other buildings, and at another moment one is fact aware that that prominent building is shorter than those other buildings, without one's being fact aware that the building has changed size. This could be the case if one's visual representations enabling fact awareness of the sizes of the building are not encoded in working memory or if those representations are never compared. Dretske could thus appeal to one of the other accounts of change blindness I discussed in the previous section. So, perhaps, we can account for change blindness on Dretske's view without concluding even that the visual representations enabling fact awareness are sparse.

The distinction between mental representations that enable verbal reports and those that do not enable verbal reports is arguably supported by folk psychology. Verbal reports express intentional states, such as thoughts and beliefs, which have intentional, or conceptual, content and mental attitudes. If one does not have the concept of a spy, one will not have the perceptual belief that there is a spy in front of one, even if one is looking at a spy. However, folk psychology is also committed to qualitative states, such as sensations, that are individuated by their qualitative characters, not by their mental attitudes or intentional content.<sup>40</sup> Such qualitative states enable one to perceptually

---

<sup>40</sup> This claim is challenged by some representationalists, e.g., Armstrong (1968) and Pitcher (1971), who claim there are no nonintentional qualitative states. However, that representationalist view is motivated by concerns about

discriminate perceptible properties, such as colors and shapes. But, although qualitative states represent the perceptible properties they enable us to discriminate, it is not clear that one must possess the concepts needed to think about such properties as colors and shapes in order to discriminate them. If one can have a visual sensation of a red square without also having an intentional state about that red square, one will see the red square without being able to report it. Likewise, if one has alternating visual sensations of the different sizes of the prominent building in the change-blindness experiment, but one does not have an intentional state, e.g., a perceptual belief, to the effect that the building is changing size, one will not report that change in size. So one's failure to report such a change does not show that one's qualitative visual states, i.e., one's visual sensations, fail to represent the different sizes of the building. So the results of change-blindness experiments do not show that the visual representations involved in visual perception are sparse in representational detail.

#### 4. Unconscious Change Perception During Change Blindness

Dretske's view and the views that change blindness is due to sparse visual representations, to vision's failure to encode all features of a scene in

---

sense data and qualia, intrinsically conscious and incorrigibly, ineffably, and exhaustively accessible properties of perception. And, as I argued in chapter 1, we can preserve the commonsense distinction between intentional states, such as thoughts and beliefs, and qualitative states, such as sensations, without committing to such properties.

working memory, and to vision's failure to compare the aspects of one's visual representations pertaining to the changing features, as well as Noë's view, all rest on the assumption that subjects who do not report the change in a changing visual scene, or who report not seeing a change in a changing visual scene, fail to see the changing features as changing.

But it could be that subjects in change-blindness experiments not only see the changing features when those features are present but also see those features as changing. As Dretske notes, the data from change-blindness experiments consist of subjects' reports about what they saw and what they did not see. But one's reports express only one's conscious mental states, i.e., those mental states one is conscious of having. So, unless one has exhaustive access to one's own mental states, one could have mental states that one is not conscious of having, i.e., states that are not conscious. It could be that subjects in change-blindness experiments do in fact see that the features are changing, but they are not conscious that they see that those features are changing. If so, those subjects will be unable to report those changes and they will be unable to report that they see those changes.

Access in the general case is inexhaustive. And there is no reason to think that one has exhaustive access to one's own mental states, including one's visual representations, or sensations. Both commonsense and experimental considerations suggest that one is often unaware of certain aspects of one's mental states. It is widely held that intentional states often occur without one's

being aware of them. For example, one's desires often guide one's behavior, even when one sincerely denies having those desires. And one often struggles and fails to recall something, e.g., someone's name, but that name suddenly pops into one's mind later, suggesting that one was wondering about it all along, even when one was not aware that one was doing so.

But cases of unconscious mental states are not limited to intentional states. In cases of subliminal perception, one is unaware that one perceives a stimulus. And studies of masked priming (Marcel, 1983; Breitmeyer et al., 2004), blindsight (Weiskrantz, 1997), and unilateral neglect (Bertelson et al., 2000), all of which rely on indirect measures of perception, also suggest that one can perceive something without being conscious that one is perceiving it. If seeing visual stimuli depends in part on one's having visual sensations that represent the properties of those stimuli, then such cases suggest that sensations sometimes occur unconsciously, i.e., without one's being conscious of them.

And we can account for change blindness in terms of the distinction between perception and conscious perception, i.e., perception of which one is not conscious. It could be that one's visual sensations of a changing visual scene do in fact represent the changing features, but one fails to access the details of the representations that are changing. In this case, one will fail to notice the change simply because one will be unaware that one has different visual sensations at those different times, even if one does in fact have such visual sensations. Without being aware that one's sensations are changing, it is

likely that one will be unaware that one is seeing a change. According to this view, change blindness results, not from a failure to see changing features, but from a failure to see them consciously.

Alternatively, it could be that one sees the changing features of a changing scene in virtue of having changing visual sensations, and one sees that those changing perceptible features are changing in virtue of having the perceptual belief that those perceptible features are changing, but one is unaware both that one's sensations are changing and that one has the perceptual belief that the features of the scene are changing. If one is unaware that one sees that the scene is changing, one will fail to report that change.

One might object that cases of so-called unconscious, subliminal perception are not cases of perception at all, so they do not provide support for the view that change blindness could occur even when one sees the change in the visual scene. Rather, one might claim, cases of so-called unconscious visual perception involve only subpersonal visual processing, not personal-level mental states such as sensations. If so, they are not cases of one's seeing something while being unaware that one is seeing it. So, one might further argue, though it could be that information about changes in visual scenes is processed at a subpersonal level during change blindness, that would not show that change blindness results from one's failing to be conscious of seeing a change that one does in fact see.

However, as I argued in chapter 1, it is unclear why one would deny that one could see a stimulus without being aware that one sees it. It is not at all obvious that folk psychology holds that all seeing is conscious. And it could be that folk psychology individuates sensations by their perceptual roles. That would allow for a folk psychological distinction between seeing and conscious seeing, i.e., seeing of which one is conscious. So it could be that change blindness occurs even when one sees the change.

The account of change blindness I am arguing for is different from Dretske's view. According to Dretske, one's reports that one sees no change, and one's failure to report changes, show that one does not see that the scene one is looking at changed. However, unless one has exhaustive access to one's visual representations and other mental states, one could see that the scene changed while failing to be aware that one sees that the scene changed. Further, this account of change blindness rests on a distinction between one's having a mental state, such as a perceptual belief or a sensation, and one's being aware that one has that mental state. Dretske's account, on the other hand, rests only on his distinction between the two ways of being visually aware of something.<sup>41</sup>

---

<sup>41</sup> There is a way in which the view I've argued for and the view that Dretske argues for could be compatible. Dretske is arguing that it could be that one consciously sees the changing features but fails to see that they have changed. If Dretske is claiming just that change blindness results from one's failing to consciously see that the features have changed, while allowing that one might unconsciously see that they have changed, perhaps our views are in fact quite close.

## 5. Experiments on Unconscious Change Detection During Change Blindness

Experiments on implicit, or unconscious, change detection suggest that one sometimes sees changes in visual scenes, even when one cannot report those changes or that one sees them (Fernandez-Duque and Thornton, 2000, 2002, 2003; Hollingworth, Williams, & Henderson, 2001; Houck and Hoffman, 1986; Laloyaux, Destrebecqz, and Cleeremans, forthcoming; Mack, 2002; Russell & Driver, 2005; Smilek et al., 2000; Thornton and Fernandez-Duque, 2000; Williams & Simons, 2000). I'll argue that such experiments show that at least some cases of change blindness are in fact due to one's failure to be conscious that one is seeing a change in a visual scene, even when one does see the change. So these experiments show that change blindness sometimes occurs even when one is in fact aware that the visual scene has changed. Dretske fails to account for such cases.

Experiments on implicit change detection during change blindness examine whether changes that subjects fail to notice, i.e., those they fail to report or those they report not seeing, result in priming effects on subjects' subsequent behavior. If changes that subjects deny seeing, or fail to report seeing, affect subjects' subsequent behavior, this suggests that subjects did in fact see those changes.

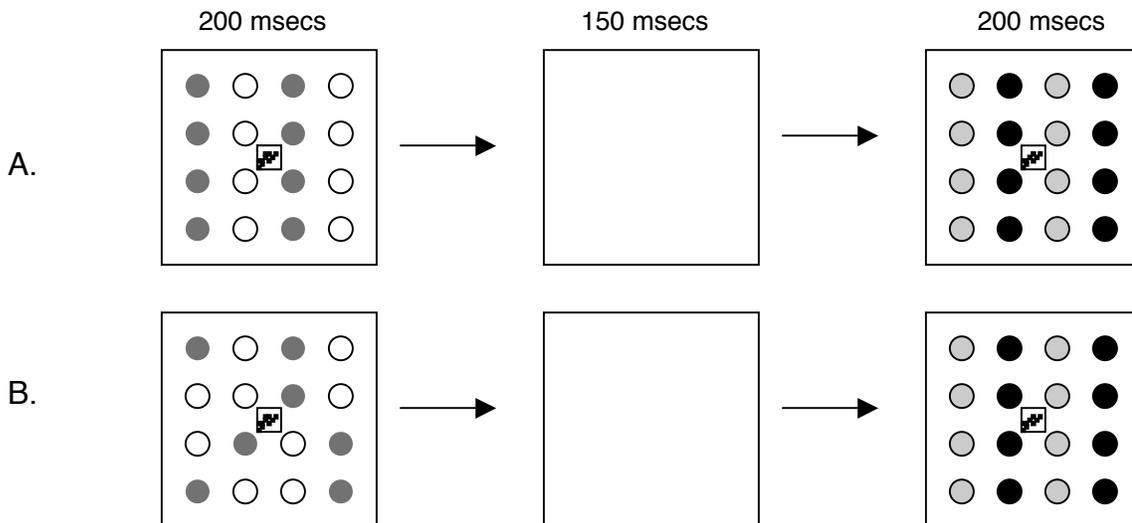
Charlotte Russell and Jon Driver (2005) found that subjects' ability to report a change in a target stimulus is influenced by the occurrence of other

changes occurring in the scene, even when subjects report seeing no changes other than those occurring to the target stimulus.

Russell and Driver instructed subjects to watch for a slight change in a small matrix of black and white pixels presented at the center of a screen. In addition to the matrix, the scene consisted of a background of sixteen dots, four across and four down. Each dot was one of two colors, e.g., red or green, and the dots could be colored so as to form columns of same-colored dots, rows of same-colored dots, or neither. For instance, when the leftmost dots are green, the dots just to the right of them are red, those directly to the right of those dots are green, and the rightmost dots are red, they appear to form two columns of green dots and two columns of red dots. Alternatively, when the uppermost dots are all red, the next four down are all green, the four directly below them are red, and the dots at the bottom are all green, they appear to form four rows of dots. And when the red and green dots are distributed randomly, they do not appear to form rows or columns.

The scene flashed for 200 msecs, then a blank screen appeared for 150 msecs, and finally a second scene appeared for 200 msecs (figure 4). After the second scene appeared, the subjects were to report as quickly as possible whether the black and white matrix in the center of the screen changed from the first scene to the second, where a change consisted of a single pixel's changing from black to white or from white to black.

Figure 1: Adapted from Russell & Driver (2005)



In trial A, the background organization is invariant from the first display to the second. In trial B, the background organization changes from the first display to the second.

Subjects reported changes in the center matrix more accurately and faster when those changes were accompanied by changes in the background organization, e.g., when the background dots changed color in a way that altered the background organization from columns to a random configuration or from a random configuration to columns.<sup>42</sup>

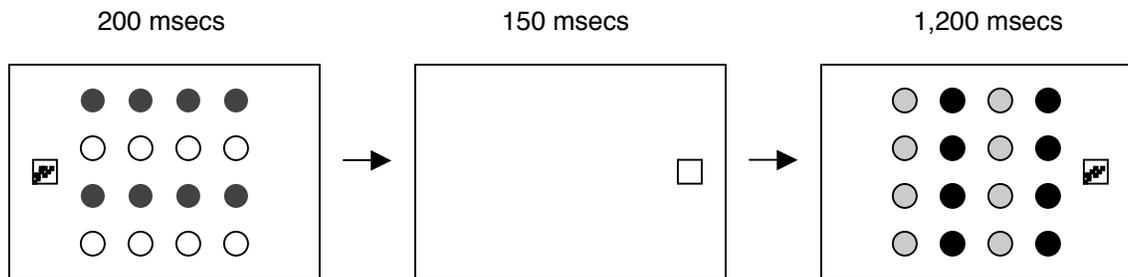
<sup>42</sup> The colors of the background dots always changed from the first to the second scene, even if the background organization did not change. For example, red and green dots forming rows could change to blue and yellow dots also forming rows. Driver et al. changed the colors of the dots because, without doing so, each change in the background organization would coincide with the change in at least some of the dots' colors. In that case, they could not determine whether the effects on one's detection of changes in the center matrix were due to the relation between those changes and changes in the background organization, as opposed to changes in the colors of the background dots.

However, although the speed and accuracy of subjects' reports about changes to the target matrix were influenced by changes in the background organization, subjects were at chance at reporting the background organizations, or even whether those background organizations changed. Though the subjects did in fact see the background organization, and they saw that it changed, they could not report that change.

In another experiment, Russell and Driver tested whether such unreported changes to the background organization of a scene affect subjects' ability to detect a small change to a target matrix when those changes occur during a saccade.

In this experiment, Russell and Driver presented subjects with an initial scene consisting of a small black and white matrix, like that in the earlier experiments, but this time the matrix was located off to the far left of the screen. The scene also contained a background consisting of 16 dots grouped by color similarity into either rows or columns, as in the previous experiments. The initial scene was presented for 200 msecs, and was followed by a screen with a small square off to the right, which served to direct the subject's gaze to that location off to the right. That screen was otherwise blank, and was presented for 150 msecs. Then a second scene consisting of a background of 16 dots grouped by color similarity into either rows or columns and a target black and white matrix positioned to the far right appeared for 1,200 msecs.

Figure 2: Adapted from Russell & Driver (2005)



As in the previous experiments, subjects were to report as quickly as possible whether the target matrix changed from the first scene to the second, where a change consisted in a single pixel's changing from white to black or from black to white. Subjects were then also asked whether the background had changed, and whether the background dots were organized into vertical columns or a random configuration.

Since the target matrix moved from the left to the right between the two scenes, subjects were required to saccade from the left to the right to perform this task.<sup>43</sup> So any changes to the target matrix or the background occurred while the subjects were saccading.<sup>44</sup>

Russell and Driver again found a congruency effect of the background organization, although the effect differed in this experiment from the effect found in the previous experiment. Unlike in the previous experiment, subjects were

<sup>43</sup> This was confirmed in a pilot study.

<sup>44</sup> Russell and Driver thus avoided the complicated use of eye-trackers involved in saccade-dependent experiments, such as those of McConkie and Zola (1979) and Grimes (1996).

neither more accurate nor faster at reporting changes to the target matrix when that change was accompanied by a change in the background organization than they were when that change was not accompanied by a change in the background organization. However, subjects were faster at reporting that the target matrix did not change when the background organization also did not change than when the background organization did change. And, again, this congruency effect did not depend on subjects' ability to report the background change or the background organization. Though subjects did not consciously see the background change, they did in fact see the change.

Pepper Williams and Daniel Simons (2000) also found priming effects of changes that subjects failed to report seeing. Williams and Simons briefly presented subjects with a novel object with multiple parts. The object then briefly disappeared and then reappeared again. When it reappeared, either it did not change or one, two, or three of its parts were changed. Subjects were to report as quickly as they could whether the object had changed from its first presentation; they were to press the S key if the object was the same as in its first presentation, and they were to press the D key if the object had changed.

Williams and Simons found that 68% of the subjects were faster at reporting that the object did not change in trials in which the object did not change than they were at reporting that the object did not change in trials in which the object did in fact change, i.e., in change-blindness trials. So subjects were slower at reporting that an object did not change when the object did in fact

change than they were at reporting that an object did not change when the object did not change. This suggests that the change in the stimulus affects subjects' response times, even though the subjects failed to report the change. Since the change affects subjects' response times, the subjects arguably saw the changes. However, since subjects in those trials reported that the stimulus had not changed, they arguably were not aware that they saw the change.

Andrew Hollingworth, Carrick Williams, and John Henderson (2001) found that subjects fixate objects that have changed longer than they fixate those same objects in control trials when they have not changed, even when subjects fail to report those changes.

Subjects viewed a line drawing of a scene, e.g., a laboratory, while their eye movements were monitored with an eye-tracking device. In some trials, after a subject fixated a particular object, e.g., a microscope, and then saccaded away from that object, the object changed, e.g., it changed into a different type of microscope. Subjects were instructed ahead of time to push a button as soon as they saw a change in the scene. In control trials, the scene remained unchanged. In all trials, the experimenters monitored subjects' eye movements, and measured how long subjects fixated the changed object when they saccaded back to it after the change.

Hollingworth et al. found that in trials in which subjects failed to report a change that occurred, i.e., in change-blindness trials, subjects fixated the changed object for a longer period of time than they fixated the unchanged object

after their initial saccade in the control trials. So the change to the visual scene affected subjects' fixation of the changed object, even when they failed to report the change. Again, this suggests that the subjects did in fact see the changes, but they were not aware that they saw the changes, i.e., they unconsciously saw the changes.

I have been arguing that the above experiments show that change-blindness subjects do in fact see changes in visual scenes, even when they cannot report those changes or that they see them. These experiments thus support the view that change blindness results, not from one's failure to see a change, but from a failure to consciously see a change.

But Stephen Mitroff, Simons, and Steven Franconeri (2002) argue that such results do not show that subjects unconsciously perceive changes in visual scenes. These experiments show that subjects unconsciously perceive changes only if they show that changes in visual scenes have subsequent effects on subjects' behavior even when subjects do not consciously perceive the changes. However, Mitroff et al. argue, it could be that subjects fail to report changes, or that subjects report that no change occurred, not because the subjects failed to consciously see the changes, but because they were not confident that they saw the changes. According to Mitroff et al., subjects could employ a conservative reporting strategy, whereby they report only those changes that they are certain they saw. Operating with such a conservative strategy, subjects will fail to report, or will deny that they saw, a change that they are not completely confident that

they saw, even if they were aware that they saw that change, i.e., even if they saw the change consciously. So, perhaps, subjects in the Russell and Driver, Williams and Simons, and Hollingworth et al. experiments consciously see the changes, but they are not highly confident that they see them. To determine whether subjects do in fact unconsciously perceive changes, one must insure that subjects employ a liberal reporting strategy, whereby they report a change whenever they have even just the slightest sense that they saw a change.

Some psychologists have attempted to insure that change-blindness subjects employ such a liberal reporting strategy by instructing those subjects to report changes whenever they think they might have seen a change, regardless of how confident they are that they saw the change. I will describe such experiments by Diego Fernandez-Duque and Ian Thornton, and I will argue that, despite objections by Mitroff et al., these experiments show that subjects sometimes see changes in visual scenes without seeing those changes consciously.

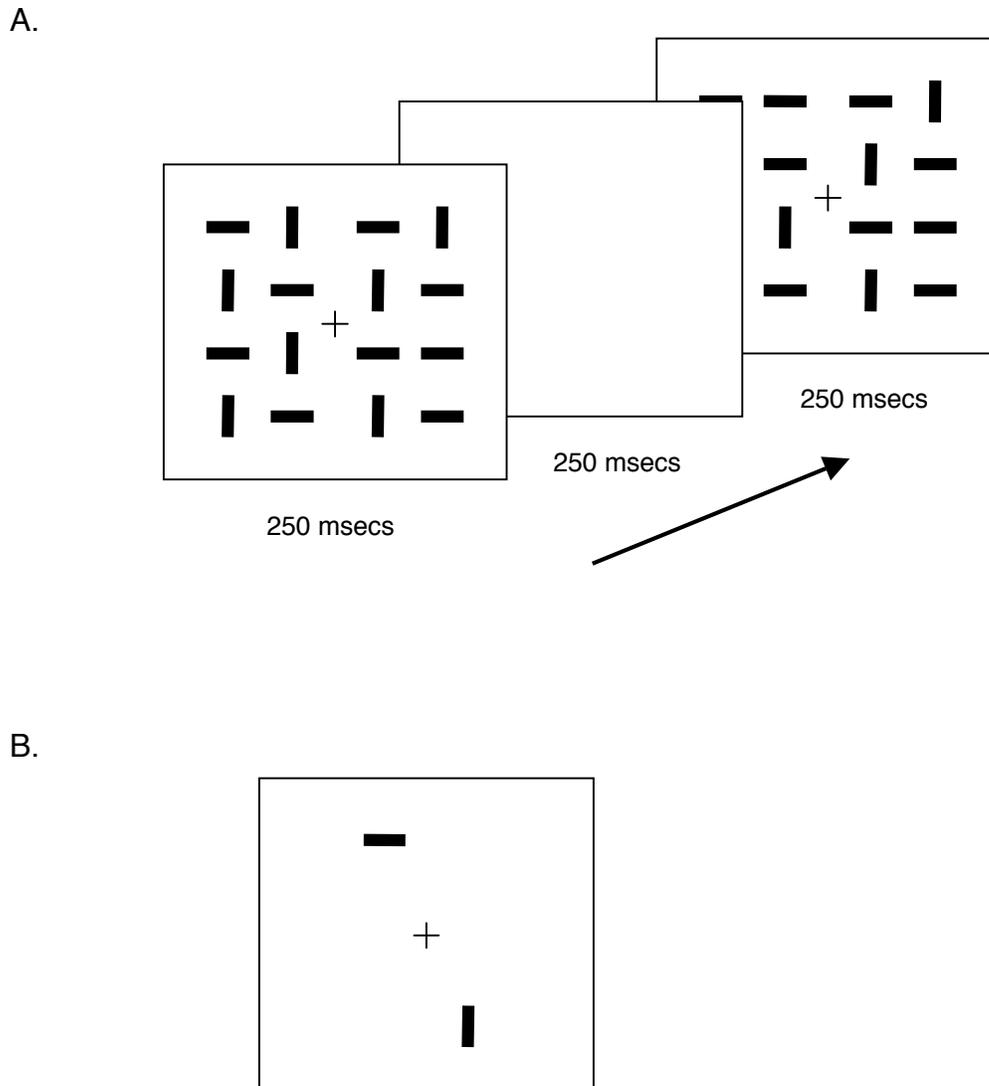
Fernandez-Duque and Thornton (2000) showed that subjects employing such a liberal reporting strategy see changes in the orientations of stimuli, even when they report seeing no such changes. Subjects were presented with a matrix of 16 black rectangles, each of which was either horizontally or vertically oriented. This matrix appeared for 250 msecs, was followed by a blank screen for 250 msecs, and then by another matrix of black rectangles for 250 msecs. The second matrix of rectangles differed from the first matrix in respect of the

orientation of one of the rectangles; that rectangle changed from horizontal to vertical or from vertical to horizontal.<sup>45</sup> When that second matrix disappeared, another screen appeared containing only two of the rectangles from the second matrix, and subjects were asked to report which of the two rectangles was most likely to have changed orientation from the first scene to the second. Subjects were then asked whether they saw a change; they were instructed to report changes if they saw a change or if they thought or sensed that they saw a change.

---

<sup>45</sup> The experiment also included catch trials in which no change occurred between the first and second matrices of rectangles.

Figure 3: Adapted from Fernandez-Duque & Thornton (2000)

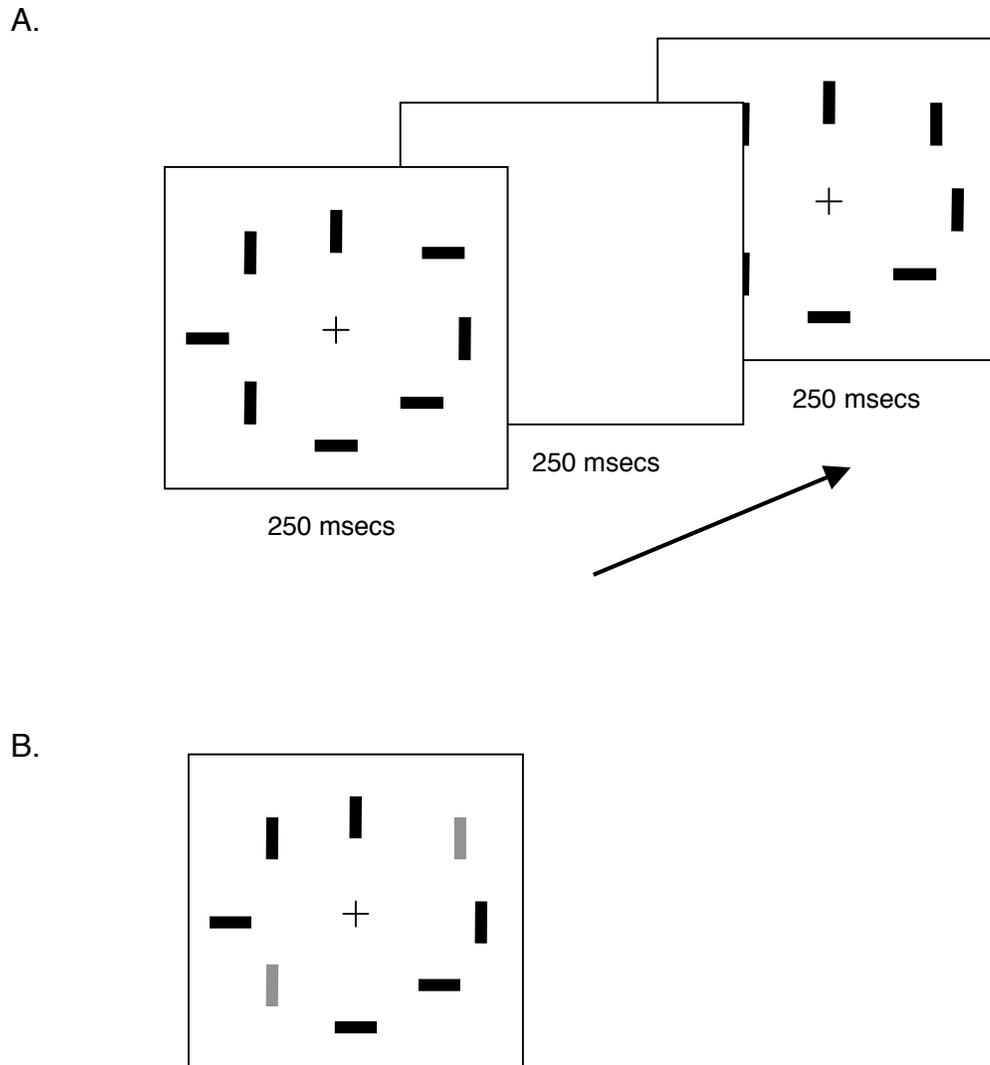


Subjects are first presented with a flicker sequence (A) in which the orientation of one of the rectangles could change after the intermittent blank screen. Then subjects are presented with a probe screen (B), and they are asked to guess which rectangle changed orientation, and to report whether they saw a change.

Fernandez-Duque and Thornton found that even when subjects reported that they saw no change, their guesses about which rectangle had changed were above chance levels. This suggests that subjects saw the change in the rectangle's orientation, even though they could not report it. And this in turn suggests that subjects fail to report the change, not because they failed to see it, but because they were unaware that they saw it.

Fernandez-Duque and Thornton confirmed these results in a variation of the experiment. They first presented subjects with eight black rectangles organized in a ring around a fixation cross such that the rectangles were equidistant from that fixation cross. Four of the rectangles were horizontal and four were vertical. This initial scene lasted for 250 msecs, was then followed by a blank screen for 250 msecs, and then by a second ring of eight rectangles for 250 msecs. In trials in which a change occurred, the second ring of rectangles differed from the first in respect of the orientation of one of the rectangles, which had changed from horizontal to vertical or from vertical to horizontal. Subjects were then presented with a scene in which two of the rectangles from the second scene were cued by changing from black to light gray. Subjects were asked to report which of those two cued rectangles had most likely changed between the first and second scenes, and then to report whether they had seen a change. Again, subjects were instructed to report a change if they saw a change or if they thought or sensed that they saw a change.

Figure 4: Adapted from Fernandez-Duque & Thornton (2000)



Subjects are first presented with a flicker sequence (A) in which the orientation of one of the rectangles could change after the intermittent blank screen. Then subjects are presented with a probe screen (B), and they are asked to guess which of the two cued rectangles changed orientation, and to report whether they saw a change.

Thornton and Fernandez-Duque again found that subjects locate the change above chance levels, even when they report that they did not see a change. Since subjects employed a liberal reporting strategy, their reports of seeing no changes arguably reflect, not a lack of confidence that they saw the changes, but that they did not consciously see the changes. These results, like those from the previous experiment, suggest that subjects saw the change in orientation, even though they were not conscious that they saw it.

However, Mitroff, Simons, and Franconeri (2002) argue that Fernandez-Duque and Thornton's experiments do not show that subjects unconsciously see the changes in orientation. Rather, Mitroff et al. argue, it could be that subjects guess above chance at the location of the change by following strategies based on what they consciously see, even though they do not see the change either consciously or unconsciously. If so, Fernandez-Duque and Thornton's experiments do not show that some cases of change blindness result from one's failure to be conscious that one is seeing a change.

According to Mitroff et al., the subjects in Fernandez-Duque and Thornton's experiments could follow an exclusion strategy, whereby they infer where a change likely occurred based on their having consciously seen that no such change occurred at another location. Subjects in Fernandez-Duque and Thornton's experiments were instructed to report which of two rectangles had changed. In trials in which a rectangle changed orientation, as opposed to catch trials, subjects were asked to choose between a rectangle that had changed and

the rectangle appearing at the location diametrically opposed to where that changed rectangle appeared. If subjects failed to see the change, whether consciously or unconsciously, but they consciously saw that the rectangle at one of the two cued locations did not change, then they could guess that it was likely that a change occurred at the other location. This strategy would of course lead to a number of false-positive responses in catch trials in which no change occurred at either location. However, it would also result in a subject's guessing above chance at the correct location of the change in trials in which a rectangle changed orientation but the subject failed to see the change at all, i.e., not even unconsciously. In such cases, subjects would report that they did not see the change, but they would correctly guess where the change had occurred. Without showing that the subjects are not following such an exclusion strategy, Mitroff et al. argue, Fernandez-Duque and Thornton fail to show that subjects see changes unconsciously, i.e., without being aware that they are seeing those changes.

Fernandez-Duque and Thornton (2003) tested whether subjects in their experiments were following such an exclusion strategy. They hypothesized that subjects using such a strategy in trials in which they reported seeing no change at all would be above chance levels at reporting the location at which no change occurred. If, in trials in which they report seeing no change, subjects guess the location of a change above chance levels because they consciously saw that no change occurred at the other cued location, subjects will report above chance levels that no change occurred at those other cued locations. If subjects' reports

of the locations at which no change occurred are not above chance, then they did not consciously see that no change occurred there. If so, subjects' above-chance guesses about the location of the change did not result from their employing an exclusion strategy; they did infer the location of change because they consciously saw that no change occurred at the other cued location.

As in their previous experiments, Fernandez-Duque and Thornton presented subjects with a ring of eight rectangles for 250 msec, followed by a blank screen for 250 msec, then another ring of rectangles for 250 msec, and finally a screen in which two of the eight rectangles were cued by changing from black to light gray. Subjects were asked to report which of the two cued rectangles was located opposite from the rectangle that they think was most likely to have changed orientation; i.e., they were asked to select the rectangle they thought most likely did not change orientation. Subjects were then asked to report whether they had seen any change in orientation.

Fernandez-Duque and Thornton found that in trials in which subjects reported seeing no change at all, they were below chance levels at selecting the rectangle located opposite from the one they thought was most likely to have changed; i.e., subjects were below chance at selecting the rectangle that had not changed orientation. Since subjects would presumably be able to select the rectangle that had not changed if they had consciously seen that it did not change, these results suggest that subjects did not consciously see that the rectangle opposite the one that changed did not change, at least not in trials in

which they report seeing no change in the scene. This in turn suggests that subjects' above-chance guessing about which rectangle changed orientation in trials in which they reported seeing no such change is not due to their using an exclusion strategy, as Mitroff et al. suggest.

In another experiment on unconscious change detection, Thornton and Fernandez-Duque (2000) tested whether changes that subjects report not seeing can nonetheless affect subjects' subsequent orientation discriminations. Specifically, they tested for a congruency priming effect, an effect whereby a change in the orientation of a stimulus affects the response times or accuracy of subsequent speeded reports about the orientation of a probe object. In many cases, a congruency between features of a previously presented stimulus and those of subsequently presented probes facilitate subjects' reports of the features of the probes, and incongruencies between features of previously presented stimuli and those of subsequently presented probes hinder subjects' reports of the features of the probe, affecting the speed or accuracy of those reports (see, e.g., Lu and Proctor, 1995; Ericksen and Ericksen, 1974; Posner, 1980; Simon and Small, 1969; Stroop, 1935). For example, Michael Posner showed that subjects are faster at reporting the appearance of a stimulus when a cue, e.g., a flash of light, appeared at the same location prior to the onset of the stimulus, whereas subjects are slower at reporting the onset of the stimulus when it is preceded by a cue appearing at a different location.

Thornton and Fernandez-Duque tested whether changes in the orientations of stimuli affect subjects' reports of the orientations of subsequently presented probes in cases in which subjects report not seeing the changes in orientation. Such a congruency effect, they argue, would show that subjects did in fact see the changes in orientation, even if they were not aware that they saw them.

As in the experiments discussed above, Thornton and Fernandez-Duque presented subjects with a ring of eight vertical or horizontal rectangles for 250 msecs, followed by a blank screen for 250 msecs, and then by another ring of eight rectangles for 250 msecs. One of the rectangles in that ring could have changed orientation from the first screen.<sup>46</sup> Finally, subjects were presented with a screen in which one of the eight rectangles was cued. Subjects were asked to quickly report the orientation of the cued rectangle by pressing one of two keys on a keyboard. They were then to press the spacebar if they thought they had seen a change in the orientation of any rectangle, or to do nothing if they thought that they had seen no change.<sup>47</sup>

There were four variations for trials in which a change occurred, valid and congruent trials, valid and incongruent trials, invalid and congruent trials, and invalid and incongruent trials. In valid trials, the probe at the end of the trial

---

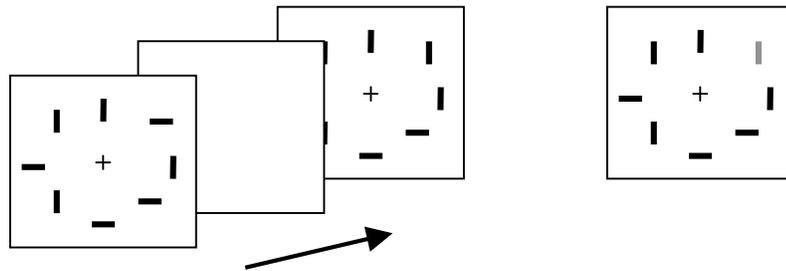
<sup>46</sup> These experiments, like the previous experiments, included catch trials, in which no change occurred between the first and second scenes.

<sup>47</sup> In a later experiment, subjects also pressed a key to report that they had not seen a change.

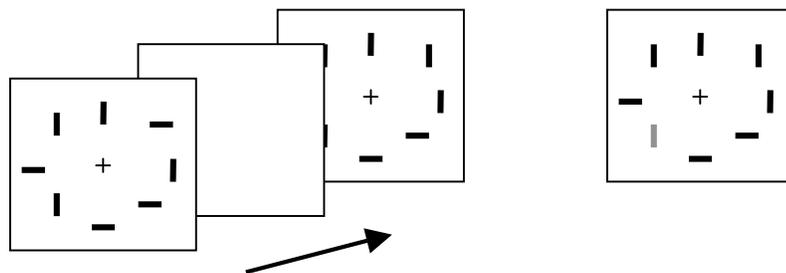
appeared at the location where a rectangle had changed orientation from the first to the second scenes. In invalid trials, the probe appeared at a different location from where the change had occurred. In congruent trials, the orientation of the probe was the same as that of the changed rectangle; e.g., if a horizontal rectangle had changed to a vertical rectangle, a vertical rectangle was cued as a probe. In incongruent trials, the orientation of the probe differed from that of the changed rectangle after the change; e.g., if a horizontal rectangle had changed to a vertical rectangle, a horizontal rectangle was cued as a probe.

Figure 5: Adapted from Thornton & Fernandez-Duque (2000)

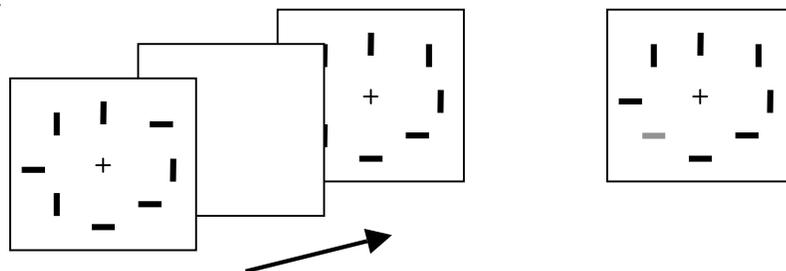
A. Valid/Congruent



B. Invalid/Congruent



C. Invalid/Incongruent



Subjects were first presented with a flicker sequence in which the orientation of one of the rectangles could change after the intermittent blank screen. In all three trials above, the rectangle at the one o'clock position changed from horizontal to vertical. Then subjects were presented with a probe screen in which one of the rectangles was cued. In valid/congruent trials, (A), the rectangle that had changed was cued. In invalid/congruent trials, (B), a rectangle of the same orientation, but at a different location from, the rectangle that changed was cued. In invalid/incongruent trials, (C), a rectangle of a different orientation from, and at a different location

from, the changed rectangle was cued. Subjects were asked to report as quickly as possible the orientation of that cued rectangle.

Thornton and Fernandez-Duque found that subjects' reports of the orientations of probes in invalid, incongruent trials were significantly less accurate than their reports of the orientations of probes in invalid, congruent trials, even in trials in which the subjects reported seeing no changes. This result, Thornton and Fernandez-Duque argue, suggests that subjects saw the changes in orientation, even when they failed to report them. On this interpretation of the data, subjects' reports of the orientations of probes are less accurate when those orientations differ from the final orientations of the changed rectangles because the final orientations of the changed rectangles primed the subjects to report those orientations, not the orientations of the probes. Vertical rectangles prime one to report vertical rectangles, whereas horizontal rectangles prime one to report horizontal rectangles. So, when one is primed to report horizontal rectangles, one will make more mistakes when reporting the orientations of vertical rectangles than when reporting horizontal ones. And when one is primed to report vertical rectangles, one will make more mistakes when reporting the orientations of horizontal rectangles than when reporting vertical ones.

It is important to note that this congruency effect is due in part to the change itself, not simply to the incongruency between the final orientation of the changed rectangle and the orientation of the probe. The probe and the changed

rectangle appeared at different locations. And some of the unchanging rectangles in the invalid, congruent trials also had orientations different from those of the probes. However, those incongruencies between the orientations of those rectangles and the orientation of the probe did not affect subjects' reports of the orientations of probes. So the best explanation of the congruency effect in invalid trials is that the final orientation of the changed rectangle primes subjects' reports. If so, the subject must have seen that rectangle change orientation. Otherwise, the orientation of the changed rectangle would not affect subjects' reports any more than the orientations of other rectangles also presented. Since this congruency effect occurs in trials in which the subjects report seeing no change at all, the subjects must have unconsciously seen the rectangle change orientation.

But, again, Mitroff et al. argue that the results of this experiment do not show that subjects unconsciously see the rectangle change orientation. Rather, they argue, the decreased accuracy of subjects' reports in invalid, incongruent trials in which subjects report seeing no change could result from subjects' awareness of an invariant spatial relation holding between the changing rectangle and the probe.

In invalid trials in Thornton and Fernandez-Duque's experiment, the probe always appeared at a location diametrically opposed to the location of the changed rectangle. It could be that subjects learn this relationship, i.e., during those trials in which they consciously see the rectangle change orientation. After

learning that the changes always occur at the location diametrically opposite from the probe, subjects could direct their attention to that location diametrically opposed to the location of the probe at the end of each trial, in which case they would attend to the changed rectangle. That in turn would increase the saliency of the orientation of that rectangle, thus creating a priming effect for subjects' subsequent reports of the orientation of the probe. If so, the congruency effect results from subjects' attending to the changed rectangle after the change, not from one's having seen the change when it occurred.

Fernandez-Duque and Thornton (2003) tested this hypothesis by running trials in which they eliminated the invariant spatial relationship between the change and the probe. Rather than presenting the probe at a location diametrically opposed to the location of the change, they systematically varied the spatial relation between the probe and the change. Fernandez-Duque and Thornton found that subjects' reports of the orientations of probes in invalid, incongruent trials were still less accurate than their reports of the orientations of probes in invalid, congruent trials, even when subjects reported seeing no change. So, Fernandez-Duque and Thornton concluded, this congruency effect does not depend on subjects' awareness of an invariant spatial relation between the probes and the changes.

However, Fernandez-Duque and Thornton note another account of the congruency effects that does not invoke unconscious perception of the change in orientation of the rectangle. After the rectangle changes orientation, the scene

contains more rectangles of one orientation than rectangles of the other orientation. In all trials, the first scene consists of a ring of four horizontal rectangles and four vertical rectangles. In trials in which one of the rectangles changes orientation, the second scene of rectangles contains five rectangles of one orientation and three of the other orientation. For example, in trials in which a horizontal rectangle changes to a vertical rectangle, the scene appearing directly after the blank screen has five vertical rectangles and three horizontal rectangles. In trials in which a vertical rectangle changes to a horizontal rectangle, that second scene contains five horizontal rectangles and three vertical ones. So it could be that the congruency effect is due to this discrepancy between the number of horizontal and vertical rectangles in the second scene. A scene with more vertical rectangles than horizontal rectangles could prime subjects to report that a subsequently presented probe is vertical. And a scene with more horizontal rectangles than vertical rectangles could prime subjects to report that a subsequently presented probe is horizontal.

To control for priming effects caused by an unequal number of vertical and horizontal rectangles in the second scene, Fernandez-Duque and Thornton ran another experiment in which each trial started with a scene consisting of an uneven number of vertical and horizontal rectangles. For example, the first scene could consist of five vertical rectangles and three horizontal rectangles. In this case, one of the vertical rectangles would change orientation in the second scene, leaving four vertical rectangles and four horizontal rectangles. If the

congruency effect in the earlier experiments resulted from an uneven number of horizontal and vertical rectangles in the second scene, the congruency effect would be eliminated in these new trials.

However, Fernandez-Duque and Thornton found that subjects' reports of the orientations of probes in invalid, incongruent trials were still less accurate than their reports of the orientations of probes in invalid, congruent trials, even in trials in which subjects reported that they saw no change. These results show that the congruency effect does not result from the presence of different numbers of horizontal and vertical rectangles in the second scene.

These results arguably support the view that subjects often see changes in visual scenes, even when they are unable to report them. Since subjects' reports about visual stimuli, such as changes, reflect only what they consciously see, change-blindness experiments show that subjects often fail to consciously see changes in visual scenes. However, since subjects' behavior, e.g., their reports of the orientation of a target object presented directly after the changing scene, is affected by changes that the subjects were not conscious of seeing, those subjects arguably saw the changes without being conscious of doing so. If so, change blindness is not a failure to see a change, it is a failure to be conscious of seeing a change that one is in fact seeing.

It is not clear how we could explain the above data on the views that change blindness results from sparse visual representations, a memory failure,

vision's failure to compare representations of the changing features before and after the change, or one's failure to see that the scene has changed.

Since subjects see the changes, they arguably see the changing features before and after the change. So their visual representations had mental qualities pertaining to those changing features, e.g., the changing orientations of the rectangles. So change blindness does not result from one's having sparse visual representations.

This also undermines much of the motivation for Noë's (2004) enactive account of change blindness, which I discussed in the previous chapter. Noë argues that change blindness shows that we see only a subset of the details of visual scenes at once. And he claims that we take visual experience to present us with so much detail, contrary to what change blindness shows, because we implicitly understand how moving will enable us to see more detail than we currently see. However, if change blindness does not show that we see very little at once, as the above experiments on unconscious change perception suggest, then we need not invoke such implicit sensorimotor understanding to explain the sense that we see a great amount of detail at once. Perhaps we seem to ourselves to see a great amount of detail at once because we do in fact see a great amount of detail at once.

Also, since the change itself produces the various priming effects discussed above, subjects arguably see that the features have changed, they do not just see the changing features without seeing them as changing. Dretske

claims that change blindness results from one's failure to see that a change has occurred, even though subjects see the changing features. So Dretske's view also fails to account for the data.

Further, since the subjects see the change in features, vision must have encoded and stored a representation of the feature of the original scene before it changed. So change blindness does not result from vision's failure to retain representations of the changing features from moment to moment.

Finally, if seeing changes requires that vision compares the representation of the scene before the change with the representation of the scene after the change, subjects' visual representations of the scene before and after the change must have been compared.

So the best explanation of the above results is that subjects see the changes in the visual scenes, but they are not conscious that they saw those changes. On this view, subjects fail to report the changes because they do not consciously see those changes, not because they do not see that those changes occurred, as Dretske argues.

#### 6. Neural Evidence for Change Perception During Change Blindness?

If subjects do in fact perceive changes when they are unable to report those changes, or even when they deny seeing those changes, then the areas of the brain underlying perception of change are presumably active during change blindness. So, one might argue, neuroscientific studies of brain activity during

change-blindness experiments could help determine whether change blindness is in fact a failure to be conscious that one sees a change, not a failure to see that change.

Diane Beck, Geraint Rees, Christopher Frith, and Nilli Lavie (2001) found that certain neural areas are activated during change blindness, i.e., when subjects fail to report a changing feature, that are not activated when subjects view a scene that does not change. This shows that, even when subjects are unable to report changes in visual scenes, their brains do register such changes. One might argue that these results support the view that subjects perceive changes in visual scenes, even when they are unable to report those changes. If so, these results support the view that change blindness is due to one's failure to be conscious that one is seeing a change.

Beck et al. used functional magnetic resonance imaging (fMRI) to monitor subjects' neural activity while the subjects were engaged in a change-detection task. The subjects were presented with a sequence of scenes starting with a scene consisting of two faces positioned on either side of a fixation cross and two strings of three letters each positioned 2.4 degrees of visual angle above and below the fixation cross. After that initial scene was briefly presented it was followed by a blank screen. Following the blank screen, another scene consisting of two strings of letters, a fixation cross, and two flanking faces appeared briefly, and was then followed by another blank screen. After subjects viewed two cycles of this sequence, they were prompted to report whether an 'X'

had appeared in the strings of letters and whether either of the two faces changed during the trial. In some trials, i.e., change trials, one of the faces changed after the intermittent blank screens; in no-change trials, neither face changed.

Beck et al. found that during trials in which a face changed but subjects failed to report that change, there was significant activation of an area of the fusiform gyrus, an area sensitive to face perception, and there was also activation of the lingual gyrus and inferior frontal gyrus. However, in trials in which neither face changed, these neural areas were not activated. Subjects' responses in both change trials and no-change trials were the same, i.e., they reported seeing no change to the faces. So, Beck et al. argue, the neural activation occurring during the change trials in which subjects failed to report the changes "... reflects stimulus-driven unconscious processing of change" (2001, p. 646).

One might argue that this supports the view that one unconsciously perceives changes. However, though it could be that the activation during change blindness reflects unconscious change perception, not all neural activity underlies psychological processing, such as perception. And it could be that the stimulus-driven activation Beck et al. found during change blindness reflects subpersonal processing of changes, not perception of changes. So these results do not by themselves provide further support for the view that subjects unconsciously see changes during change-blindness experiments.

To determine whether the neural activity occurring during change blindness does in fact underlie change perception, one could monitor neural activity during a change-blindness experiment in which subjects exhibit priming effects from changes they fail to report, such as those found by Fernandez-Duque and Thornton (2000, 2003), Hollingworth et al. (2001), Russell and Driver (2005), and Williams and Simons (2000). In addition to monitoring neural activity, one could test whether activation of those neural areas is required for change perception by applying transcranial magnetic stimulation (TMS) to those areas while subjects are engaged in a similar test for unconscious change perception. TMS temporarily deactivates the neural areas to which it is applied. So, if applying TMS to those areas reduces or eliminates the priming effects of changes one fails to report, this suggests that those areas are in fact required for unconscious change perception.

However, other results in the Beck et al. study might pose a problem for the view that the stimulus-driven activation occurring during change blindness is identical with, or even underlies, unconscious change perception. If so, the activations of the fusiform gyrus, lingual gyrus, and inferior frontal gyrus that Beck et al. found do not support the view that subjects unconsciously see changes during change blindness.

Beck et al. found that the neural areas activated when subjects report the changes, so when they consciously perceive those changes, are separate from the areas activated during change blindness; there is no overlap between the

areas activated during conscious change perception and the areas activated during change blindness. Beck et al. claim that the lack of overlap between those activations shows that the results of experiments showing priming effects of changes during change blindness are due to unconscious change perception, not to low-confidence conscious change perception, as some (e.g., Mitroff et al., 2002) argue. If the priming effects were due to low-confidence conscious change perception, Beck et al. argue, the neural processes causing them would presumably involve the same, but weaker, neural processes as the high-confidence conscious change perception underlying subjects' successful reports of changes.

However, one might argue that the lack of overlap between areas activated during change blindness and areas activated during conscious change perception suggests that the areas activated during change blindness are not identical with unconscious change perception. If unconscious change perception occurs, it arguably involves the process in virtue of which one perceives changes but not the processes in virtue of which one is conscious of perceiving changes. And conscious change perception arguably involves both the process in virtue of which one perceives changes and the process in virtue of which one is conscious of perceiving changes. If so, conscious and unconscious change perception involve common processes, in virtue of which one perceives changes in both cases. Since the Beck et al. experiments suggest that there is no overlap between the neural processes activated during conscious change perception and

those activated during change blindness, one might argue, the processes occurring during change blindness are not identical with, and are not even involved in, unconscious change perception. If so, perhaps the Beck et al. findings do not support the view that subjects see changes even when they fail to see them consciously.

Still, it could be that conscious change perception does in fact involve those areas activated during change blindness, but the processes Beck et al. found to be activated during conscious change perception suppress or mask those processes. On this view, the activation occurring during change blindness also occurs during conscious change perception, but the fMRI fails to detect it.

Other results might support this hypothesis. Luiz Pessoa and Leslie Ungerleider (2004) found that the neural areas activated during conscious change perception are virtually identical with the areas activated during false alarms, i.e., cases in which subjects reported seeing changes during trials in which no changes occurred. So it could be that the activation Beck et al. detected during conscious change perception does not include processing involved in change perception, just processing involved in one's being conscious of oneself as seeing changes and in one's reporting changes.<sup>48</sup> Since conscious change perception, as opposed to cases of false alarms, does in fact involve the

---

<sup>48</sup> Although it could be that false alarms are caused by illusory states of change perception, or that the processes in virtue of which one falsely reports a change cause illusory states of change perception, we need only conclude that this activation common to conscious change perception and false alarms underlies the processes in virtue of which one is conscious of oneself as seeing a change and in virtue of which one reports changes.

perception of changes, it arguably involves processing that was not detected by fMRI.

One could test whether conscious perception of changes requires activation of those areas Beck et al. found to be activated during change blindness by applying TMS to those areas in subjects engaged in a change-detection task. If such TMS lowers or eliminates subjects' ability to report changes, compared with subjects' success when TMS is not applied to those areas, then those areas are in fact involved in conscious change perception, even if they are not detected by fMRI during conscious change perception.

But other experiments pose another problem for the view that the areas activated during change blindness underlie change perception. Fernandez-Duque, Giordana Grossi, Thornton, and Helen Neville (2003) measured event-related potentials (ERPs) of activation during both change blindness and conscious change perception, and they found that those areas activated during change blindness respond to the changes in the visual scenes after a longer delay than the areas activated during conscious change perception respond to the onset of such changes. Again, if the activation that occurs during change blindness underlies change perception, then the processes it underlies are arguably involved in conscious change perception as well. Again, this is because conscious change perception arguably involves both the process of perceiving changes and the process in virtue of which one is conscious of perceiving changes. Further, the processes involved in change perception are presumably

causally antecedent to those in virtue of which one is conscious of perceiving a change; one would not be conscious of perceiving a change before one sees a change. Since the areas activated during change blindness respond to changes more slowly than the areas activated during conscious change perception, activation of those areas does not cause the activation that occurs during conscious change perception. So, one might argue, those areas activated during change blindness do not underlie perception of change.<sup>49</sup>

Nevertheless, this does not pose a problem for the view that subjects do in fact see changes while being unable to report them. It could be that the activation that occurs during change blindness and the activation that occurs during conscious change perception are both caused by the same processes, and that those earlier processes are identical with change perception. And it could be that those earlier processes are suppressed or masked by the processes detected by fMRI. There is visual processing that occurs earlier than the processing found during these experiments. So even if the areas activated during change blindness are not activated during conscious change perception, that does not show that unconscious change perception does not also occur.

---

<sup>49</sup> Fernandez-Duque et al. claim these results show there are separate processes underlying conscious and unconscious change detection. However, that conflicts with the simpler view that conscious change perception involves the process in virtue of which one sees change, which is also involved in unconscious change perception, and the process in virtue of which one is conscious of seeing that change.

So, whereas the neuroscientific experiments on change blindness do not support the view that change blindness is a failure to be conscious that one is seeing a change, they do not pose a problem for that view either.

However, despite the inconclusiveness of the neuroscientific studies of brain activity during change blindness, the psychological experiments revealing priming effects of changes during change blindness strongly suggest that subjects do see changes in visual scenes, even when they fail to report them or when they report not seeing them. So we can best explain change blindness in terms of the view that subjects are sometimes unaware that they see changes in visual scenes that they nonetheless do in fact see.

### 7. Homomorphism Theory and Change Blindness

If change blindness occurs when subjects see changes and changing features without being aware that they see them, then change blindness occurs even when subjects have both visual representations that represent those changing features and visual representations that represent those changes as such. So we must explain the nature of such visual representations in a way that allows for the distinction between one's having such representations and one's being conscious of them.

It is widely held that intentional states, such as beliefs, often occur without one's being conscious of them.<sup>50</sup> This is because it is widely held that intentional states are to be individuated by their mental attitudes and conceptual content, both of which are amenable to functional explanation in terms of their typical causes and effects. Since one's states can play the functional roles of intentional states without one's being conscious of those states, intentional states can occur unconsciously. And, since perceptual beliefs are intentional states, they arguably sometimes occur without one's being conscious of them. If so, one's perceptual belief that a visual scene is changing could occur unconsciously.

However, seeing also involves qualitative states, or sensations. And qualitative states are individuated, not only in terms of their functional roles--i.e., their roles in mediating between sensory inputs, other mental states, e.g., perceptual beliefs, and behavioral outputs--but also in terms of their qualitative characters. Since it is unclear why one would fail to consciously see a change as such if one consciously sees the changing features of a scene, the best explanation of change blindness is that subjects do not consciously see the changing features, or at least that they do not consciously see the changing feature either before or after the change. If so, the view that change blindness results from one's failure to be conscious that one is seeing a change rests on the view that qualitative states, such as visual sensations, can occur without one's being conscious of them.

---

<sup>50</sup> But see Galen Strawson (1994) for an argument that intentional states do not occur unconsciously.

The homomorphism theory of sensing I argued for in Chapter 1 explains the nature of visual sensations in a way that is compatible with this account of change blindness, since homomorphism theory allows for a distinction between one's having sensations and one's being aware of those sensations (Rosenthal, 1991, 2005).

According to homomorphism theory, visual sensations represent visual stimuli and scenes in virtue of having mental qualities that are analogous in a specific way to the visible properties of those stimuli and scenes. Specifically, mental qualities represent perceptible properties in respect of homomorphisms between families of mental qualities and families of perceptible properties. For example, visual sensations represent perceptible colors in virtue of having mental qualities, i.e., colors\*, that resemble and differ from each other in ways parallel to the ways perceptible colors resemble and differ from each other. Just as the color red resembles orange more than green, red\*, the mental quality of sensations of red, resembles orange\* more than green\*. Likewise, visual sensations of shape have mental qualities, shapes\*, that resemble and differ from each other in ways parallel to the ways visible shapes resemble and differ from each other. Just as perceptible squares resemble perceptible rectangles more than perceptible triangles, square\* resembles rectangular\* more than triangular\*. And the same account explains how visual sensations represent size, orientation, and location.

This view best explains how we see various perceptible properties, and how we perceptually discriminate those properties. According to homomorphism theory, we see, e.g., squares as more similar to rectangles than triangles because visual sensations of squares are more similar to those of rectangles than those of triangles. Homomorphism theory holds that mental qualities are theoretical posits posited to explain how we see perceptible properties, such as the shapes, sizes, orientations, and locations of objects.

Homomorphism theory also accounts for the introspectible qualitative character of conscious visual sensations. What it's like for one to consciously see a square is more similar to what it's like for one to consciously see a rectangle than what it's like for one to consciously see a triangle. This is because when one introspects the visual sensation one has when seeing a square, one is aware that that sensation itself resembles and differs from other visual sensations of shape in ways parallel to the ways the perceptible shapes they enable one to see resemble and differ from each other.

But, according to homomorphism theory, since visual sensations are individuated in terms of the properties in virtue of which they enable us to see perceptible properties, not in terms of how we are conscious of those sensations, one can have, e.g., a visual sensation of a square without being aware that one is having it, so without there being anything it's like for one to see that square. Homomorphism theory thus accounts for how one can unconsciously see features of a visual scene.

This in turn provides an account of how visual sensations represent stimuli and scenes independently of one's being aware of those sensations. So one could see the changing features of a visual scene in virtue of having changing visual sensations of those features, but without being aware of the changes in one's sensations. If so, one would fail to report the change. Also, if one sometimes sees something without being conscious of seeing something, it could be that one sees the change in the scene as such, but one fails to be conscious of seeing it. In this case, one would fail to report seeing the change, and one would deny seeing the change if asked.

This view of change blindness rests of course on a distinction between one's seeing something, e.g., a change, and one's being aware that one is seeing it. So the view is committed to a distinction between the mental processes in virtue of which one sees something and the mental processes in virtue of which one is aware of seeing something. Accordingly, when we consciously see something, we are aware of seeing it in virtue of having a higher-order mental state that represents the first-order mental state enabling us to see. We can account for how we are aware of the first-order states in virtue of which we see stimuli in terms of a higher-order theory of consciousness, such as Rosenthal's (1997, 2005) higher-order-thought model of consciousness. According to Rosenthal, one consciously sees something when and only when one has a seemingly noninferential thought to the effect that one is seeing that thing. Accordingly, if one fails to have a higher-order thought to the effect that

one is seeing a change in a visual scene that one does in fact see, one will fail to see that change consciously, so one will fail to report that change.<sup>51</sup>

Though I have argued that the results of experiments on change blindness are best explained in terms of one's failure to be aware that one sees a change, it could be that some cases of change blindness are due to other causes, such as vision's failure to represent the changing features, to encode them in working memory, or to compare them after the change. Homomorphism theory is compatible with all of these other accounts of change blindness.

For example, homomorphism theory is compatible with the view that visual representations are sparse, not highly detailed. A sparse visual sensation, on this view, is one that has mental qualities pertaining to only a small subset of the visible properties of the visual scene. If one had such a sparse sensation, one would see only those perceptible properties that one's sensation represented. If the sensation lacked the mental qualities pertaining to the changing details of a scene, one would fail to see those details as well as the change.

---

<sup>51</sup> There are a number of factors that could contribute to a higher-order thought's failing to represent one's seeing a change. For example, it could be that one's consecutive higher-order thoughts fail to represent one as seeing the features that in fact change, even though one does in fact see those features. Alternatively, it could be that one's higher-order thoughts do represent one as seeing those features at the times when one sees them, but one fails to retain the earlier higher-order thought long enough for one to notice that one has seen a feature change. Or, perhaps, though one does retain that higher-order thought, one simply fails to draw an inference between that higher-order thought and the current higher-order thought that represents one as seeing the new, changed feature.

Homomorphism theory is also compatible with the view that change blindness results from vision's failure to retain representations of visible features from moment to moment, since homomorphism theory does not hold that visual sensations and their mental qualities always persist. It could be that when one is presented with the picture of a city's skyline, one's visual sensation does have mental qualities pertaining to all of the buildings, but that sensation is not stored in working memory. In that case, when one forms a new sensation with mental qualities pertaining to the properties of the changed scene, i.e., the scene in which one of those buildings is 25% larger than in the first scene, one will fail to notice that the scene changed.

Further, homomorphism theory is compatible with the view that visual representations are highly detailed but the mechanism that compares successive representations of scenes does not compare all of the details of those representations. If vision fails to compare the mental qualities of one's sensations pertaining to the changing building before and after the change, then one will fail to see the change, even though one's sensations represented the features of that changing building before and after the change.

Finally, homomorphism theory is compatible with Dretske's view of change blindness. If the mental qualities of one's sensations pertaining to the changing features of a scene also change, but one does not form the perceptual belief that the scene is changing, then one will not report the change, and one will deny seeing the change if asked. According to homomorphism theory, mental

qualities of visual sensations are posited to explain how one visually discriminates visible features. But one could do so without being able to report the features one is discriminating. As Dretske claims, verbal reporting requires intentional states. So, if one lacked the belief that the visual scene was changing, one would not report that it was changing, even if one's sensations were changing.

## Chapter 4: Feature Binding and Multiple-Object Tracking

### 1. Introduction

We discriminate stimuli on the basis of not only their individual perceptible properties, such as their colors or shapes, but also their combinations of perceptible properties. For example, one can discriminate a scene consisting of a red square next to a green triangle and a scene consisting of a green square next to a red triangle, even though both scenes contain the same colors and shapes. To discriminate these scenes one must see the difference in how those colors and shapes are combined.

Sensing such feature conjunctions is not limited to the visual case. One senses a combination of perceptible properties when one feels something cold, hard, round, and smooth, when one tastes something both sweet and spicy, and when one has a pain that is both dull and throbbing. Any theory of sensing must account for our sensing such feature conjunctions.

Austen Clark (2000, 2004) argues that we sense feature conjunctions in virtue of sensing distinct perceptible properties at the same location. On this view, one sees a red square in virtue of seeing red and seeing a square at the same place, and one feels something both smooth and cold in virtue of feeling smoothness and coldness at the same place. According to Clark, this way of explaining how one senses feature conjunctions requires a special treatment of how one senses the locations of stimuli. Specifically, Clark argues that whereas

homomorphism theory, the view of sensing I advocate in earlier chapters, adequately explains how one senses other perceptible properties, e.g., colors, textures, and pains, we need a separate account of how one senses the locations of stimuli. Clark offers his Feature Placing view as the best account.

Zenon Pylyshyn (2003) offers an alternative to Clark's view.<sup>52</sup> According to Pylyshyn, one senses a combination of perceptible properties by sensing them as properties of the same object, not as properties instantiated at the same location. On Pylyshyn's view, seeing a red square, e.g., rests on two distinct operations. First, vision picks out an object without representing any of that object's properties, including its location. And, once vision has picked out an object in this primitive, nonrepresentational way, it forms representations of that object's color and shape. Accordingly, one sees an object as both red and square in virtue of forming a visual representation of red and a visual representation of a square in connection with the same primitive, nonrepresentational access to the object.<sup>53</sup>

---

<sup>52</sup> Jonathan Cohen (2004) and Mohan Matthen (2004) argue for views similar to Pylyshyn's.

<sup>53</sup> Such access to objects is nonrepresentational only in the sense that it is supposed to occur independently of representing any of an object's properties, e.g., its location, color, or shape. One could of course argue that such access is in fact representational. It is often claimed that linguistic demonstratives, such as 'this' and 'that', represent objects without representing any of their properties. Accordingly, Pylyshyn's claim is not that vision picks out objects nonrepresentationally, just that it picks them out independently of representing any of their properties. Whether we call such access to objects representational or nonrepresentational is merely a verbal issue.

Pylyshyn's view rests on arguments that nonrepresentational access to objects is needed to explain the effects of so-called object-based attention and to explain how we keep track of a number of moving visual stimuli at once, as revealed by his multiple-object-tracking experiments.

If Pylyshyn's object-based view of sensing feature conjunctions is correct, Clark's location-based view is false. If so, we do not need a special account of sensory localization, as Clark argues.

I'll argue that Pylyshyn's view of sensing feature conjunctions is unmotivated and problematic, and thus fails to undermine Clark's location-based view. But I'll then argue that Clark's view is superfluous and therefore fails to undermine homomorphism theory's account of how one senses the locations of objects. In so arguing, I'll argue that we can best explain how one senses combinations of perceptible properties in terms of the view that distinct mental qualities, e.g., those in virtue of which we see color and shape, are interdependent, and that this view is a consequence of homomorphism theory.

## 2. Homomorphism Theory and the Many-Properties Problem

According to homomorphism theory, one senses perceptible properties, such as colors and shapes, in virtue of having sensory states with mental qualities that represent those perceptible properties. On this view, mental qualities represent perceptible properties in virtue of resembling and differing from other mental qualities in ways parallel to the ways the perceptible properties

resemble and differ from each other (Rosenthal, 1991, 2005; Meehan, 2002). For instance, the sensation one has when one sees an ellipse is more similar to the sensation one has when one sees a circle than it is to the sensation one has when one sees a triangle, just as ellipses are more similar to circles than they are to triangles. This is because the sensations have mental qualities, shapes\*, that resemble and differ from each other in ways parallel to the ways visible shapes resemble and differ from each other. And it is because such mental qualities resemble and differ in ways parallel to the ways perceptible properties resemble and differ from each other that we can sense the similarities and differences among the perceptible properties. This homomorphism theory, I argue, applies to cases of sensing all sensible properties, e.g., colors, sounds, textures, bodily stimulation, and all sensible spatial properties, e.g., shapes, sizes, orientations, and locations.

But Clark argues that homomorphism theory fails to explain how one senses feature conjunctions. Clark invokes Frank Jackson's (1977) so-called many-properties problem to motivate his view of sensing feature conjunctions and to argue against homomorphism theory.

One can distinguish the following two scenes:

- a) A red square next to a green triangle.
- b) A green square next to a red triangle.

Since both scenes contain the same colors and shapes, seeing the difference between these scenes requires more than just seeing their colors and shapes. One can distinguish these two scenes because one sees the differences in the ways the colors and shapes are combined. When looking at scene (a), one sees a combination of red and square and a combination of green and triangular; when looking at scene (b), one sees a combination of green and square and a combination of red and triangular. Jackson's many-properties problem is the problem of explaining how one sees such combinations of properties.

According to Clark, one sees a combination of properties in virtue of sensing those properties at the same location. When one sees scene (a), one sees red and square at the same location and green and triangular at another location, and when one sees scene (b), one sees green and square at the same location and red and triangular at another location.

But, Clark argues, this location-based solution to the many-properties problem is unavailable to views, such as homomorphism theory, that account for one's sensing the locations of stimuli in terms of one's having sensations with mental qualities that represent those locations (2000, p. 68). According to such views, one sees something off to the left, e.g., in virtue of having a left\* sensation, or a sensation off-to-the-left\*.<sup>54</sup> Accordingly, if one sees a red square off to the left, one has a red\*, square\* sensation off-to-the-left\*. But this view, Clark argues, fails to account for one's capacity to discriminate between cases

---

<sup>54</sup> I suffix an asterisk, i.e., "\*", to a predicate to signify reference to a mental quality, as opposed to the perceptible property that mental quality represents.

such as (a) and (b), since, according to this view, seeing both could involve sensations with exactly the same mental qualities.

Suppose in scene (a) the red square is at location L1 and the green triangle is at L2, and in scene (b) the green square is at L1 and the red triangle is at L2. If homomorphism theory is true, Clark claims, one's sensation of scene (a) will have the mental qualities red\*, green\*, square\*, triangular\*, L1\*, and L2\*, but one's sensation of scene (b) will also have the mental qualities red\*, green\*, square\*, triangular\*, L1\*, and L2\*. So to explain how one discriminates between scenes (a) and (b), homomorphism theory must explain how vision binds particular colors\* and shapes\* to particular locations\*, e.g., to bind red\* and square\* to L1\* and green\* and triangular\* to L2\*. But, according to Clark, homomorphism theory does not explain how distinct mental qualities are bound to each other. So, he claims, though we sense other perceptible properties, e.g., colors, in virtue of having sensations with mental qualities that represent those properties, we do not sense where objects are located in virtue of having sensations with mental qualities that represent the locations of those objects.

Rather, Clark claims we can best explain how one discriminates cases such as (a) and (b) in terms of the view that sensing involves two distinct mechanisms, one in virtue of which we sense the locations of stimuli and the other in virtue of which we sense perceptible properties at those locations (2000, p. 74; 2004, p. 450). According to Clark, we sense various properties, such as colors and textures, in virtue of having sensations with mental qualities, e.g.,

colors\* and textures\*, as characterized by homomorphism theory, but we sense where those properties are located in virtue of mechanisms called sensory names.

On this view, which Clark calls Feature Placing, one sees a red patch off to the left when a sensory name picks out a region off to the left in one's field of view, the space in which one sees stimuli at a given moment, and one has a red\* sensation in connection with that sensory name. Accordingly, one sees a red square in virtue of having instantiations of the mental qualities red\* and square\* both in connection with the same sensory name.

According to Clark, when one sees a red square next to a green triangle, one has instantiations of red\* and square\* in connection with the same sensory name and instantiations of green\* and triangular\* in connection with another sensory name. And when one sees a green square next to a red triangle, one has instantiations of green\* and square\* in connection with the same sensory name and instantiations of red\* and triangular\* in connection with another sensory name. Feature Placing can thus solve the many-properties problem.

### 3. Experimental Support for Clark's Location-Based View

In support of his location-based solution to the many-properties problem, Clark cites work on the binding problem in neuroscience, the problem of explaining how the brain gives rise to unified sensations of feature conjunctions, given that it represents distinct visible properties, e.g., color, shape, orientation,

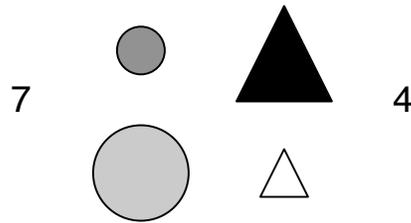
size, and motion, in different parts of visual cortex. According to Clark, properties of these separate neural representations are identical with the mental qualities of sensations (2000, p. 44). So, Clark assumes, a solution to the binding problem constitutes a solution to the many-properties problem.

Clark focuses primarily on Anne Treisman's influential work on the binding problem (2000, p. 46; 2004, p. 449-451), especially her experiments on illusory conjunctions and conjunctive-feature searches (see Treisman, 1999 for a review). I'll focus on illusory conjunctions here.

In some of the earliest experiments on illusory-conjunctions, Treisman and Hilary Schmidt (1982) showed that subjects sometimes accurately report the visible properties of stimuli present while inaccurately reporting combinations of those properties. They take these results to reveal a failure of the mechanism responsible for binding separate visual representations of distinct perceptible properties.

In one experiment, subjects were briefly presented with a number of colored shapes flanked by two black numerals; e.g., subjects were presented with a small blue circle, a larger yellow circle, a small pink triangle, and a larger brown triangle flanked on either side by a black '7' and '4' (figure 1). Before the scene appeared, subjects were instructed to give priority to remembering the numerals for a subsequent memory task.

Figure 1: Illusory Conjunction Paradigm



Color Key:  Blue  Brown  Yellow  Pink

As soon as the scene disappeared, subjects were asked to report the numerals. Also, a marker appeared where one of the colored shapes had been, and subjects were asked to report as many of that object's properties as they could.

Treisman and Schmidt found that subjects reported a significant number of illusory conjunctions, i.e., conjunctions of properties that were not present in the same stimuli. An example of such an illusory-conjunction report is when a subject reports a small pink circle when presented with small blue circle, a larger yellow circle, a small pink triangle, and a larger brown triangle. Though some illusory-conjunction reports included properties not presented in the display, e.g., red or square, most of the illusory conjunctions subjects reported combined properties that had been present in the display.

Treisman and Schmidt take such illusory-conjunction experiments to show that one can accurately see the perceptible properties present while misperceiving their combinations. So they conclude that illusory-conjunction

reports reveal an error of whatever mechanism binds separate neural representations of distinct perceptible properties.

According to Treisman and Schmidt, these studies suggest that binding separate neural representations of distinct visible properties requires focal attention, a limited-capacity processing mechanism that enables vision to process information about certain stimuli at the expense of processing information about other stimuli seen at the same time. They base this conclusion on their finding that subjects are more likely to report illusory conjunctions when they are allocating such attention to the flanking numerals for the subsequent memory task rather than attending to the colored shapes.

The view that one must attend to a stimulus to see feature conjunctions is further supported by studies showing that a patient with Balint's Syndrome, a visual-attention deficit caused by bilateral parietal lesions, frequently reports illusory conjunctions when presented with multiple colored shapes (Friedman-Hill et al., 1995; Robertson et al., 1997; Cohen and Rafal, 1991).

However, evidence that focal attention is responsible for seeing feature conjunctions does not, by itself, support Clark's location-based solution to the many-properties problem. But several experiments suggest that focal attention is allocated to locations in one's field of view, facilitating processing of features instantiated at those locations (Posner, 1980; Treisman & Gelade, 1980; Treisman, 1988). If focal attention is required for seeing feature conjunctions, and it operates on locations, then perhaps seeing feature conjunctions is location

based, as Clark claims. According to this view, one sees combinations of perceptible properties when focal attention is allocated to a particular location and it determines that those perceptible properties are present there.

Michael Posner (1980) showed subjects respond faster to objects appearing at previously cued locations, e.g., if a light briefly flashes there preceding the target. According to Posner, subjects are faster at responding to such an object because the cue attracts focal attention to where the object will appear, and attention then persists at that location long enough to coincide with the object's appearance.

Treisman (1988) found that pre-cueing an object's location facilitates subjects' reports of conjunctions of that object's properties more than it facilitates reports of single properties of the object. According to Treisman, this occurs because allocating attention to a location enables one to see feature conjunctions there, whereas attention plays no such role in one's seeing single features.<sup>55</sup>

Since it seems that spatially allocated focal attention is required for one to see feature conjunctions, Treisman concludes that neural representations of distinct perceptible properties are bound when focal attention determines that the perceptible properties they represent are present at the same location.

According to Treisman, focal attention does this by determining that separate

---

<sup>55</sup> The claim that attention plays a role in seeing feature conjunctions but not single features is supported by Treisman's work with Garry Gelade on conjunctive-feature searches (Treisman & Gelade, 1980).

neural representations of those perceptible properties represent those properties as being at the same location. So Treisman offers a location-based solution to the binding problem. Since Clark thinks the properties of those separate neural representations are the mental qualities in virtue of which one senses perceptible properties, he adopts a location-based view to solve the many-properties problem.<sup>56</sup>

Treisman (1999) also cites a number of other sources in support of her location-based view of binding. For one, Mary Jo Nissen (1985) found a statistical dependency between one's ability to report an object's shape and one's ability to report its location, and also between one's ability to report an object's color and one's ability to report its location. But Nissen found that no such correlation exists between one's ability to report an object's shape and color. According to Treisman, this suggests that vision represents properties, such as color and shape, along with their locations in one's visual field, and that it represents those distinct properties separately. Further, if vision represents both colors and shapes along with their locations, then it can bind separate representations of colors and shapes by correlating them with regard to the locations at which they represent those colors and shapes.

And Treisman and Gelade (1980) found a significant statistical correlation between subjects' correctly reporting feature conjunctions and their correctly

---

<sup>56</sup> Clark claims his view does not rest on all of the details of Treisman's view (2004, p. 446). But the details to which Clark is noncommittal are not relevant to this discussion.

reporting the locations of those features. Subjects were briefly presented with two rows of colored letters, each of which was a pink 'O' or a blue 'X', with the exception of the target letter, which was either a pink 'X' or a blue 'O'. After the brief presentation, subjects were to identify which type of target was present and where it was located.

Treisman and Gelade found that when subjects reported the target's location incorrectly they were at chance at identifying which kind of target, a pink 'X' or a blue 'O', it was. But in a similar task in which targets were distinguished from distractors by only a single feature, rather than by a feature conjunction, subjects could identify targets even when they could not locate them. Treisman and Gelade take this to support their location-based solution to the binding problem. They claim there is a strong correlation between one's ability to identify a feature conjunction and one's ability to locate that feature conjunction because seeing feature conjunctions depends on vision's determining that distinct features are present at the same location.

Finally, Asher Cohen and Richard Ivry (1989) showed that subjects are more likely to report illusory conjunctions involving features of objects that are located close to each other than features of objects farther apart. Again, this suggests a correlation between vision's binding representations of distinct features and the location at which vision represents those distinct features.<sup>57</sup>

---

<sup>57</sup> See Clark (2004) for more support for the location-based view.

#### 4. Pylyshyn's Object-Based View

Zenon Pylyshyn (2003) offers an alternative view, according to which we see feature conjunctions when vision represents distinct features as features of the same object, not as features at the same location. According to Pylyshyn, the experiments frequently cited in support of the location-based view equally support this object-based alternative.

... in all studies that examine the mislocation of properties, as for example in the case of conjunction illusions, location and object identity (i.e., which object it is) are confounded, since the objects have fixed locations: in this case being a particular object O is indistinguishable from being at location X. Because of this, the findings are equally compatible with the view that individual objects as such are detected first, before any of their properties (including their locations) are encoded. (2003, p. 180; emphasis in original)

Pylyshyn further argues that we do in fact see conjunctions of features, e.g., color and shape, in virtue of vision's representing them as features of the same object, not as features at the same location. This object-based view has two main sources. One source is the growing experimental research suggesting that the kind of focal attention involved in seeing feature conjunctions is allocated to objects, not locations, so seeing feature conjunctions rests on preattentive

access to objects in virtue of which vision allocates attention to those objects.

The second source of support for the object-based view of seeing feature conjunctions comes from Pylyshyn's experiments on our ability to simultaneously keep track of several moving visual stimuli.

I'll argue that both considerations fail to show that seeing feature conjunctions depends on vision's representing distinct features as features of the same object. So, I'll argue, Pylyshyn fails to undermine Clark's location-based account of sensing feature conjunctions.

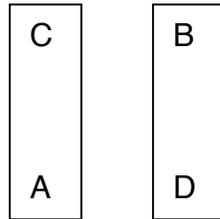
#### 5. Pylyshyn's Object-Based View: Object-Based Attention

Along with Treisman, Pylyshyn thinks focal attention is responsible for binding separate representations of distinct perceptible properties. But Pylyshyn argues, against Treisman, that focal attention is allocated to objects, not locations. If focal attention is allocated to objects, not locations, and if seeing feature conjunctions requires focal attention, vision must access objects in a way that enables attention to be allocated to those objects prior to seeing feature conjunctions, so seeing feature conjunctions rests on this preattentive access to objects. So, Pylyshyn thinks, vision binds distinct representations of distinct perceptible properties when focal attention determines that those representations represent properties of the same object, not properties at the same location.

In support of this view, Pylyshyn cites growing experimental support for the claim that focal attention operates on objects, not locations (2003, ch. 4; also see Scholl, 2001 for a review).

Robert Egly, Jon Driver, and Robert Rafal (1994) showed that though cueing a location in one's field of view speeds one's reports of stimuli at nearby locations, as Posner showed, one's responses are even faster when the cue and the stimulus appear within the same object. Egly et al. presented subjects with two parallel, rectangular bars. One end of one of the bars was then cued by a brief color change. After the cue disappeared, the luminance of one of the ends of the two bars changed. The luminance change could occur at the cued end of the pre-cued bar, at the non-cued end of that bar, or at either end of the bar that was not pre-cued. Subjects were to report whether that luminance change occurred at the cued location, i.e., whether the trial was valid.

Figure 2: Adapted from Egly et al. (1994)



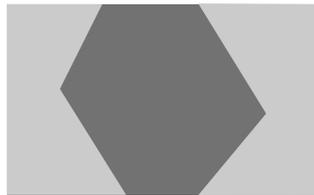
The cued end is marked here with a 'C'. A trial in which the luminance of any other end changes is an invalid trial. Subjects are faster at reporting an invalid trial when the luminance changes at the end marked with the 'A' than when the luminance changes at the ends marked with the 'B' or 'D', even though the ends marked with the 'A' and the 'B' are equidistant from the end marked with the 'C', i.e., the cued end.

Egly et al. found that subjects are faster at reporting invalid trials, i.e., those in which the luminance change occurred elsewhere from the cue, when the luminance change occurred within the cued bar rather than within the non-cued bar. This same-object effect occurs even when the luminance change in the non-cued bar and the luminance change at the non-cued end of the cued bar are equidistant from the location of the cue. If the speed of such reports depends on the allocation of attention, and attention is first allocated to the cued end, then these results show that shifting attention within an attended object is easier than shifting attention between objects. According to Egly et al., this supports the view that attention is allocated not to locations but to objects.

Gordon Baylis and Driver (1993) showed subjects are faster at reporting the relative locations of two features, e.g., vertices, when those features occur

within a single object than when they occur in different objects. They presented subjects with displays similar to that in figure 3 and instructed them to report which vertex, the left or right, is higher. Before the trials began, one group of subjects was instructed to attend to the red region and another group was instructed to attend to the green regions. Both groups of subjects were presented with the same displays. Baylis and Driver found subjects attending to the red region are faster at reporting the relative heights of the vertices than subjects who were instructed to attend to the green regions.<sup>58</sup>

Figure 3: Adapted from Baylis and Driver (1993)



Color Key:  Green  Red

According to Baylis and Driver, we cannot account for these results in terms of the view that attention is allocated to locations because the locations of the vertices are the same in both cases; subjects attending to the red region and those attending to the green regions were presented with the same displays.

---

<sup>58</sup> In other trials the colors were switched, so parts that are red in figure 3 were green and the parts that are green in figure 3 were red. The colors of the regions were irrelevant to the results.

However, we can account for these results in terms of the view that attention is allocated to objects. It could be that subjects instructed to attend to the red region see that region as a single object, so they see the vertices as features of that particular object, whereas subjects instructed to attend to the green regions see the vertices as features of two distinct, green objects. Accordingly, the subjects attending to the red region are faster at reporting the relative heights of the vertices because focal attention facilitates processing of information about objects, i.e., focal attention is object based. When one sees the vertices as features of the same object, attention need not shift between objects for one to determine the relative heights of the vertices. On the other hand, when one sees the vertices as features of different objects, one's attention must shift between the two objects to determine which vertex is higher. If attention is allocated to objects, shifting one's attention between objects would result in a processing cost, which could explain why subjects attending to the green areas are slower at reporting which vertex is higher.

If focal attention is allocated to objects, not locations, then perhaps binding is object based, as Pylyshyn argues.<sup>59</sup>

---

<sup>59</sup> Pylyshyn's object-based view of binding is compatible with the homomorphism theory of sensing. It could be that one sees a green, square object and a red, triangular object in virtue of vision's forming a green\*, square\* representation of one object, and a red\*, triangular\* representation of another object.

## 6. Objections to Pylyshyn: Binding Without Attention

Pylyshyn's argument that binding is object based, like Treisman's argument that binding is location based, thus rests on the view that focal attention is required for binding separate representations of distinct perceptible properties, e.g., color and shape. And this view derives from experiments such as Treisman's on illusory-conjunction reports, which are supposed to show that a limited-capacity processing mechanism, i.e., focal attention, is required for one to see feature conjunctions.

But illusory-conjunction experiments do not show that attention is required for seeing feature conjunctions. False reports of feature conjunctions are reports of feature conjunctions nonetheless. For example, when a subject wrongly reports a small, blue triangle when presented with only a large red triangle, a small red triangle, a small blue square, and a large blue square, the subject does in fact report a conjunction of size, shape, and color. So, if Treisman is right that such illusory-conjunction reports result from a failure to allocate attention to the target stimulus, attention is not required for seeing feature conjunctions.<sup>60</sup> Perhaps attention facilitates accurate sensing of feature conjunctions, but it is not required for sensing those feature conjunctions.

---

<sup>60</sup> Of course, it could be that one does not fail to attend to the colored shapes, but spreads one's attention across the entire scene. If so, such limited attention to the colored shapes could explain one's illusory-conjunction reports. Still, these experiments do not show this is the case.

Further, illusory-conjunction experiments show only that, under certain conditions, subjects inaccurately report feature conjunctions; they do not show that subjects inaccurately see those feature conjunctions. Reports of visual stimuli sometimes fail to reflect what one in fact sees. For example, cases of subliminal perception and experiments on masked priming (Marcel, 1983), implicit change detection (Fernandez-Duque & Thornton, 1999; Hollingworth, Williams, & Henderson, 2001), inattention blindness (Mack & Rock, 1998), unilateral neglect (Bertelson et al., 2000), and blindsight (Humphrey, 1983; Weiskrantz, 1997) show that one can see stimuli while being unable to report them. So it could be that subjects accurately see the feature conjunctions, even when they inaccurately report them.

And experiments by Russell and Driver (2005)<sup>61</sup> and Michael Houck and James Hoffman (1986) show that one can in fact see feature conjunctions one cannot report. This shows that what is required for one to report feature conjunctions is distinct from what is required for one to see them. So it could be that subjects in Treisman and Schmidt's illusory-conjunction experiments accurately see feature conjunctions but inaccurately report them. In this case, the experiments reveal a failure of the mechanism in virtue of which one accurately reports feature conjunctions, not a failure of a mechanism in virtue of which one sees them. Even if illusory-conjunction experiments showed that

---

<sup>61</sup> I discussed these experiments earlier in chapter 3.

illusory-conjunction reports result from a failure of attention, they would not show that attention is required for seeing feature conjunctions.

Driver et al. and Houck and Hoffman further argue that their experiments show one can see feature conjunctions, not only without being able to report them, but without attending to the objects with those features. If so, attention is not required for seeing feature conjunctions, as Treisman and Pylyshyn claim.

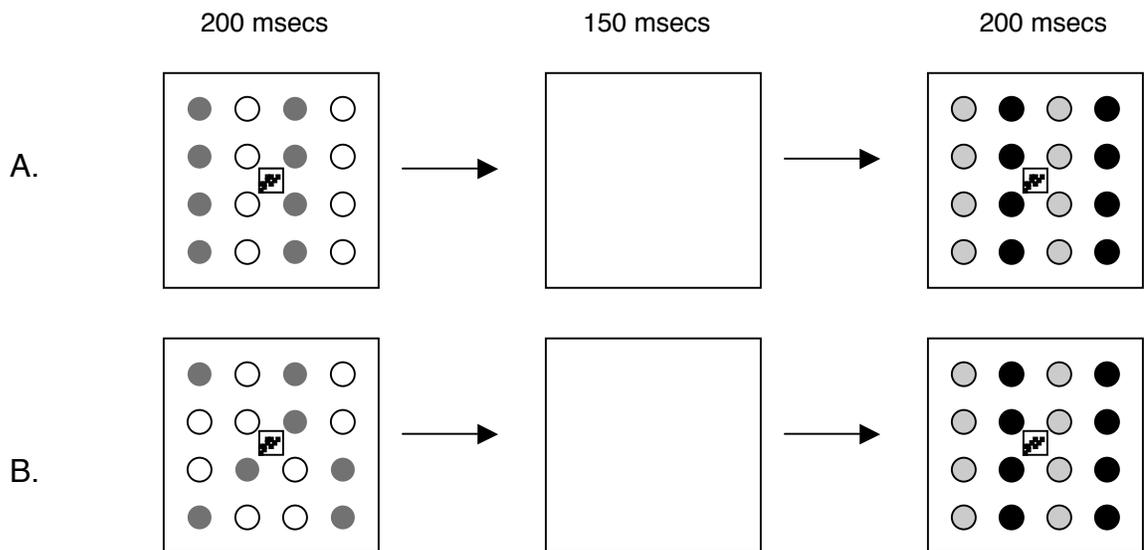
However, I'll argue that the Russell and Driver and Houck and Hoffman experiments fail to show that attention is not required for seeing feature conjunctions, though they do show that one can see feature conjunctions without being able to report them. So, I'll argue, these experiments do not show that one can see feature conjunctions without attention, but they do undermine Treisman and Schmidt's claim that illusory-conjunction experiments show focal attention is required for one to see feature conjunctions.

Russell and Driver instructed subjects to watch for a slight change in a scene consisting of a small matrix of black and white pixels that was surrounded by sixteen dots, four across and four down. Each dot was one of two colors, e.g., red or green, and the dots could be colored so as to form columns of same-colored dots, rows of same-colored dots, or neither. For instance, when the leftmost dots are green, the dots just to the right of them are red, those directly to the right of those dots are green, and the rightmost dots are red, they appear to form two columns of green dots and two columns of red dots. Alternatively, when the uppermost dots are all red, the next four down are all green, the four directly

below them are red, and the dots at the bottom are all green, they appear to form four rows of dots. And when the red and green dots are distributed randomly, they do not appear to form rows or columns.

The scene flashed for 200 msecs, then a blank screen appeared for 150 msecs, and finally a second scene appeared for 200 msecs (figure 4). After the second scene appeared, the subjects were to report as quickly as possible whether the black and white matrix in the center of the screen changed from the first scene to the second, where a change consisted of a single pixel's changing from black to white or from white to black.

Figure 4: Adapted from Russell and Driver (2005)



In trial A, the background organization is invariant from the first display to the second. In trial B, the background organization changes from the first display to the second.

Subjects reported changes in the center matrix faster or more accurately when such changes were accompanied by changes in the background organization, e.g., when the dots changed color in a way that changed the background organization from rows to columns or from rows to a random configuration.<sup>62</sup> And when no change occurred in the center matrix, subjects were faster or more accurate at reporting that there was no change when the background organization also remained unchanged.

However, subjects were unable to report the background organizations. They could not report whether they had seen, e.g., rows or columns or randomly arranged dots. And, according to Russell and Driver, subjects don't attend to the background dots; the task requires them to attend only to the black and white matrix. Nevertheless, the speed and accuracy of subjects' reports are affected by changes in the background organization. This shows that the subjects did in fact see the background organization, even if they did not attend to it and could not report it. Further, seeing the background organization depends on one's seeing the colored dots as parts of larger colored objects with certain orientations, i.e., rows or columns, and that organization is a function of the

---

<sup>62</sup> The colors of the background dots always changed from the first to the second scene, even if the background organization did not change. For example, red and green dots forming rows could change to blue and yellow dots also forming rows. Russell and Driver changed the colors of the dots because, without doing so, each change in the background organization would coincide with the change in at least some of the dots' colors. In that case, they could not determine whether the effects on one's detection of changes in the center matrix were due to the relation between those changes and changes in the background organization, as opposed to changes in the colors of the background dots.

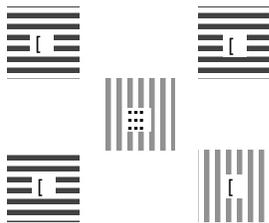
colors and relative locations of the dots. So to see the rows and columns one must see conjunctions of color, shape, and orientation. So, Driver et al. conclude, subjects see feature conjunctions without being able to report them and without attending to the objects that have those features.

Houck and Hoffman (1986) showed that the McCollough Effect, a visual aftereffect caused by certain conjunctions of color and orientation, can occur without one's being able to report those feature conjunctions and without one's attending to the stimuli causing the aftereffect. The McCollough Effect occurs when one views two alternating gratings each composed of differently oriented and differently colored lines, such as a grating of horizontal green lines alternating with a grating of vertical red lines. After extensive exposure to such alternating gratings, the subject is presented with a grating of either horizontal or vertical white lines of the same spatial frequency as those presented earlier. Subjects experience an aftereffect such that the white lines appear the complementary color of that of the previously presented lines of the same orientation; i.e., horizontal white lines appear pink and vertical white lines appear green.

Houck and Hoffman tested whether this aftereffect depends on one's attending to the gratings in the adaptation phase. They hypothesized that if seeing conjunctions of color and orientation requires attention, the aftereffects caused by unattended alternating gratings will be weaker than those caused by attended gratings.

To test this hypothesis, Houck and Hoffman presented subjects with a central grating surrounded by either four or eight other gratings. Each grating was composed of either green or magenta horizontal or vertical lines. They alternated each grating with a grating of complementary color and orientation every 5 seconds, so each grating could produce its own aftereffect.

Figure 5: Adapted from Houck & Hoffman (1986)



During the adaptation phase, the five gratings alternate with gratings of complimentary color and orientation every 5 seconds. During this phase, subjects performing the central task monitor the matrix of dots to detect the disappearance of a dot in the middle row, subjects performing the peripheral task monitor the brackets to detect a change of one of these left brackets to a right bracket, and subjects performing the dual task monitor both the matrix of dots and the brackets.

To control where subjects attended, Houck and Hoffman assigned them one of three change-detection tasks to perform during this adaptation phase, a central task, a peripheral task, or a dual task. Every 500 ms, a left bracket, i.e., '[', appeared in the center of each peripheral grating for 300 ms. At some point, a right bracket, i.e., ']', appeared in one of these gratings instead of the left bracket.

Also, three rows of three small dots were presented in the central grating. Every 100 ms one of the nine dots briefly disappeared, with a dot in the middle row disappearing only 10% of the time. Subjects performing the central task were to indicate when a dot from the middle row disappeared. Subjects performing the peripheral task were to indicate when and where a right bracket appeared in a peripheral grating. And subjects in the dual task were to simultaneously perform both peripheral and central tasks. Houck and Hoffman reasoned that subjects would allocate attention in accord with their designated task, i.e., subjects performing the central task would attend only to the central grating, subjects performing the peripheral task would attend only to the peripheral gratings, and subjects performing the dual task would divide attention among the peripheral gratings and the central grating.

Houck and Hoffman found subjects performing the dual task were significantly less accurate in reporting the presence of targets, i.e., right brackets or extinguished dots in the middle row, than subjects performing only the central task or only the peripheral task. According to Houck and Hoffman, this shows that subjects allocated their attention in accordance with their assigned tasks; subjects performing the dual task performed worse than subjects performing a single task, according to Houck and Hoffman, because the former divided their attention between the peripheral and central gratings, whereas the latter focused their attention on their designated target gratings.

Houck and Hoffman hypothesized that if attention is required for one to see conjunctions of orientation and color, unattended gratings will produce weaker aftereffects. If so, aftereffects from the peripheral gratings would be less robust than those from the central gratings for subjects performing the central task and aftereffects from the central grating would be less robust than those from the peripheral gratings for subjects performing the peripheral task.

However, Houck and Hoffman found McCollough aftereffects were as robust for unattended gratings as for attended gratings. For example, subjects performing the central task claimed aftereffects caused by the peripheral gratings were as robust as those caused by the central grating. Since the McCollough effect depends on one's seeing conjunctions of orientation and color, Houck and Hoffman claim their results suggest that subjects saw such feature conjunctions even when they were not attending to the objects with those orientations and colors.

They also found that when subjects were asked to describe the gratings at each position directly following the trials, they reported the combinations of orientation and color at chance levels. Houck and Hoffman take this to show that subjects did not consciously see conjunctions of the colors and orientations of the gratings.

If Russell and Driver and Houck and Hoffman are right, seeing conjunctions of color and orientation does not require attention and one can see such conjunctions even when one is unable to report them.

But Russell and Driver and Houck and Hoffman assume that subjects' inability to accurately report conjunctions of features of the background stimuli and to accurately report other stimuli, such as the brackets in the Houck and Hoffman experiments, shows that the subjects are not attending to those stimuli.<sup>63</sup> However, it could be that one can attend to something without being able to report it. If so, subjects in these experiments could be attending to the background stimuli even though they are unable to report those stimuli. If subjects do attend to the background stimuli, their attention could enable them to see the feature conjunctions. So Russell and Driver and Houck and Hoffman fail to show that attention is not required for one to see feature conjunctions.

In fact, recent experiments on a blindsight subject suggest that one can attend to stimuli without being able to report them (Kentridge, Heywood, and Weiskrantz, 2004).

Blindsight is a phenomenon affecting patients with lesions to primary visual cortex that prevent them from being able to report and, in most cases, act on stimuli presented in a certain region of their field of view. However, blindsighted subjects are able to guess above chance levels at the nature of

---

<sup>63</sup> Both Russell and Driver and Houck and Hoffman provide their subjects with tasks designed to focus attentional resources on one part of the screen, thus ignoring the rest of the screen. So, one might argue, there is good reason, apart from their reports, to think that subjects do not attend to the background in those experiments. However, the assumption that such tasks prevent one from attending elsewhere on the screen rests on prior experiments on attention, e.g., Treisman's and Posner's, that rely on subjects' reports. Since those experiments do not test for unreportable attentional effects, it is not clear that the tasks in Houck and Hoffman's and Russell and Driver's experiments control for subjects' attending elsewhere, e.g., to the background stimuli.

certain stimuli presented in their blind field. For example, they are above chance at guessing the orientation of a line presented in their blind field, even though they cannot report the orientation of the line, or even that the line is present.

Kentridge et al. tested blindsight subject GY to determine whether an attentional cue facilitates his guessing the orientation of a line presented in his blind field. First, an arrow pointing in one of two directions appeared at a fixation point in the center of a screen. When the arrow disappeared, a small, black, horizontal or vertical bar appeared in GY's blind field either at the location in which the arrow had been pointing or in another location. Since this bar was located in GY's blind field, he could not report it. When the bar appeared, a 200 ms tone prompted GY to guess the bar's orientation.

Kentridge et al. found GY was significantly faster and no less accurate at guessing the orientation of bars located where the arrow had been pointing than he was at guessing the orientation of bars located where the arrow had not been pointing. They claim this shows that the arrow effectively cued GY's focal attention. And since GY reported seeing nothing in these trials, this suggests he attended to stimuli he could not report.

If one can attend to stimuli one cannot report, as the Kentridge et al. experiment suggests, the Driver et al. and Houck and Hoffman experiments do not show that one sees feature conjunctions without attention.

Still, the Kentridge et al. experiment casts doubt on Treisman's conclusion that illusory conjunctions result from a failure of attention. Kentridge et al. show

that one can attend to a stimulus that one cannot report. So reporting a stimulus requires some mechanism over and above attention. Treisman's conclusion that attention is required for seeing feature conjunctions rests on subjects' false reports of feature conjunctions. But it could be that those false reports result from a malfunction of the mechanism required for reporting stimuli, not from a failure of attention. Therefore, Treisman's illusory-conjunction experiments do not show that seeing feature conjunctions requires attention. And, again, if attention isn't required for one to see feature conjunctions, then binding could be location based even if attention is object based, as Pylyshyn argues.

In any case, even if seeing feature conjunctions, and not merely reporting them, does require attention, as both Treisman and Pylyshyn claim, there could be both location-based and object-based mechanisms of attention. So it could be that location-based attention, not object-based attention, is required for seeing feature conjunctions. If so, the occurrence of object-based attention does not undermine Treisman's and Clark's location-based views of binding.

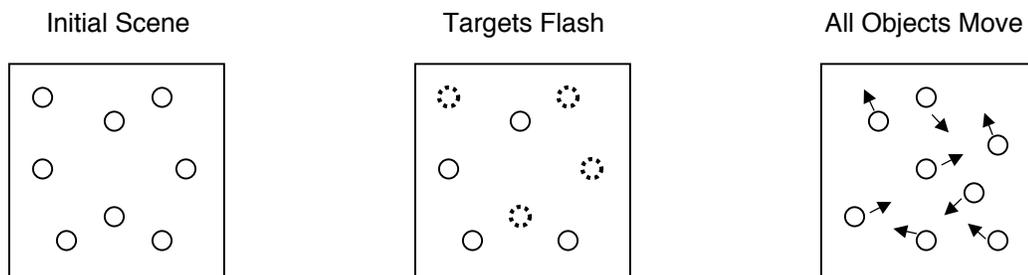
## 7. Multiple-Object Tracking

Nevertheless, Pylyshyn has another argument for his object-based view of sensing feature conjunctions. According to Pylyshyn, our ability to simultaneously keep track of multiple moving visual stimuli shows that vision accesses objects independent of representing any of their properties, including their locations. If vision accesses objects without representing any of their

properties, it could use that access to bind distinct representations of distinct perceptible properties, such as color and shape. If so, we see conjunctions of perceptible properties in virtue of representing them as properties of the same object, not as properties at the same location.

Pylyshyn and Ron Storm (1988) showed that subjects can keep track of four or five out of ten identical, moving stimuli at a time. Subjects are presented with ten identical, stationary stimuli, e.g., ten small, blue circles, four of which briefly flash to designate them as targets. All ten circles then begin moving in unpredictable but continuous pathways, and subjects are to keep track of the four targets. After a while, the objects stop moving, one of them is designated with a marker, and the subject is to report whether that object was a target. Pylyshyn and Storm found subjects reported correctly 86% of the time, suggesting that they can successfully track four or five such targets at a time.

Figure 6: Adapted from Pylyshyn & Storm (1988)<sup>64</sup>



<sup>64</sup> A demonstration of MOT is available at Brian Scholl's website: <http://pantheon.yale.edu/~bs265/demos/mot.html#mot>.

According to Pylyshyn, such multiple-object-tracking (MOT) experiments show that vision accesses objects independent of representing any of their properties. If vision did not access objects this way, Pylyshyn argues, we could not successfully track multiple objects at the same time.

The objects in the MOT task are identical with respect to all of their visible properties except for their locations. Each object is a blue circle of the same diameter, e.g., and is distinguished only by its location on the screen. So one's visual representations of the objects differ only with respect to where they represent the objects as being located; just as each object has a unique location, each visual representation of an object represents it at a unique location. So, for every target object one sees, one has a visual representation that uniquely picks it out in virtue of representing the object's unique location.

But once the targets and distractors all begin to move, they all change location, and the representations that uniquely picked out targets a moment earlier now fail to pick them out. Instead, new representations representing the targets at their new locations uniquely pick out the targets.

To see a target at that moment as the same as a target one saw a moment earlier vision must somehow correlate the representation that uniquely picks out that target with the representation that uniquely picked it out a moment earlier. But there is nothing that the new representation of that target has in common with the old representation of the target that it doesn't also have in common with the other representations of the targets and distractors. The old

target representation, the new target representation, one's representations of the other targets, and one's representations of the distractors all represent objects with the same color, size, and shape and different locations. So it might seem that there is no way vision can determine which new representation to correlate with the old target representation to enable tracking. Without an explanation of how vision correlates the new representation of that target with the old representation of that target, we cannot explain MOT in terms of representations of objects' properties. So, Pylyshyn claims, we track multiple objects in virtue of some nonrepresentational access to targets.

If this argument is sound, it shows that Clark's view of sensing fails to account for MOT. According to Clark, one sees where a small, blue circle is located in virtue of a sensory name's firing and thus picking out the circle's unique location. But when one is looking at a moving object, different sensory names will pick it out at different moments. So, to explain how one tracks an object in a MOT task, Clark must explain how vision determines which sensory names pick out the same object at different moments. Without such an explanation, Clark's view fails to explain MOT.

#### 8. What About Represented Proximity?

One might argue that vision could correlate visual representations of the same target at consecutive moments on the basis of the proximity of the locations at which those representations represent that target at those moments.

At time T1, one sees a target at location L1 in virtue of having a visual representation of it at L1. At time T2, the target is at location L2, and one has a representation of it at L2. If L2 is closer to L1 than any location of any other target or distractor is at T2, vision could correlate one's representations of that particular target at T1 and T2 in virtue of the closeness of the locations at which those representations represent the target at those two moments. Since one's representation of the target at T2 represents features of the object at a location closest to L1, the location of the target a moment ago, vision correlates one's representation of that target with one's representation of it a moment earlier, and one is able to track the target from T1 to T2.

But Pylyshyn argues that this model fails to account for MOT. His argument is based on a computer simulation of MOT in which he and Storm modeled a mechanism that correlates visual representations of the same targets at consecutive moments in terms of the proximity at which they represent the targets at those moments.

Pylyshyn and Storm designed their simulation on the assumption that vision would use focal attention to compare and correlate visual representations at consecutive moments. Because of this, the speed at which the simulation compared representations was constrained by the fastest recorded speed of focal attention.<sup>65</sup> The commands for the simulation were as follows:

---

<sup>65</sup> Posner (1978) recorded focal attention at 4 ms/degree.

- 1) While the targets are visually distinct, scan attention to each target and encode its location on a list. Then, when targets begin to move, do steps 2-6.
- 2) For  $n = 1$  to 4, check the  $n$ th position in the list and retrieve the location  $Loc(n)$  listed there.
- 3) Scan attention to  $Loc(n)$ . Find the closest object to  $Loc(n)$ .
- 4) Update the  $n$ th position on the list with the actual location of the object found in (3). This becomes the new value of  $Loc(n)$ .
- 5) Move attention to the location encoded in the next list position,  $Loc(n + 1)$ .
- 6) Repeat from (2) until elements stop moving. Go to each  $Loc(n)$  in turn and report elements located there. (Pylyshyn, 2003, pp. 224-5)

Pylyshyn and Storm found that this simulation tracked targets at an 8% success rate, not the 86% success rate of human subjects performing the same MOT task. In another simulation, they found that if the model also uses information about the direction in which the targets are moving to predict where they will be a moment later, it still tracks the objects with less than 20% accuracy.<sup>66</sup> Since the model was significantly less

---

<sup>66</sup> In another simulation, Pylyshyn and Storm determined this model could track the objects at a 39.8% success rate if the tracking mechanism accidentally recovered some targets it had previously lost track of.

successful than the human subjects, Pylyshyn and Storm concluded that we do not track multiple objects by a serial mechanism, i.e., focal attention, that compares and correlates representations of targets' locations on the basis of represented proximity or direction of movement.

Rather, Pylyshyn claims that subjects track multiple objects in virtue of four or five mechanisms each dedicated to tracking its own target. According to this view, these mechanisms, called visual indexes, are mental analogues of the linguistic demonstratives 'that' and 'this'; they pick out objects independent of representing any of their properties, including their locations (2003, p. 254). Pylyshyn claims visual indexes are causally activated by targets when those targets flash at the beginning of a trial, and they remain attached to the targets as they move, thus enabling one to track the targets through those movements.

#### 9. Visual Indexes Are Unmotivated and Problematic

Pylyshyn's argument that we see conjunctions of distinct features, e.g., color and shape, in virtue of vision's representing them as properties of the same object rests on the above argument that vision accesses objects without representing any of their properties. If vision accesses objects in this primitive, basic way, Pylyshyn assumes, then presumably it uses such access to form bound representations of those objects' properties. But, I'll argue, Pylyshyn fails

to show that such access is required for MOT, so his account does not undermine Clark's location-based view of sensing feature conjunctions.

According to Pylyshyn, the simulation shows that MOT requires primitive, nonrepresentational access to objects by showing vision could not track multiple objects by correlating representations of targets' locations from moment to moment. Pylyshyn and Storm simulated a single mechanism, i.e., focal attention, that correlates representations of objects' locations serially, and they found that such a mechanism is too slow to track the targets with the same success as human subjects performing the same tracking task. Since this simulation failed, Pylyshyn and Storm conclude that we do not track multiple objects by a single visual mechanism that correlates representations of objects' locations in this way. And they further conclude that vision tracks multiple objects via multiple mechanisms operating in parallel and picking out objects independent of representing their locations.

But the view that we track objects by correlating representations of their locations is compatible with Pylyshyn's view that we track multiple objects via multiple tracking mechanisms operating in parallel. It could be that four or five mechanisms, each dedicated to a particular target object, enable us to track multiple objects by correlating representations of their targets' locations. If so, correlating those representations would not be constrained by the speed of a single tracking mechanism such as focal attention, as Pylyshyn and Storm assume. Such mechanisms would perform MOT significantly better than the

mechanism Pylyshyn and Storm simulated, especially if those mechanisms also exploited information about the direction in which a target is moving. So Pylyshyn and Storm's simulation fails to show that MOT requires primitive, nonrepresentational access to objects.

It is crucial to the model of MOT that I have suggested that a tracking mechanism's being dedicated to a target object does not rest on the kind of primitive, nonrepresentational access to an object Pylyshyn argues for.

On the view that vision tracks objects in virtue of correlating representations of those objects, vision forms representations of those objects prior to tracking them. Vision could represent an object as, e.g., a small blue disk at location L1, another object as a small blue disk at L2, another as a small blue disk at L3, and yet another as a small blue disk at L4. When the objects flash in the beginning of a MOT trial, vision could assign each of the four tracking mechanisms to a target object picked out by its unique location. Tracking mechanism A could be assigned to the flashing blue disk at L1, which A picks out by way of vision's unique representation of that object. Tracking mechanism B could be assigned to the flashing blue disk at L2, and so on for mechanisms C and D. So the dedication of each tracking mechanism to its target rests on vision's representing that target by way of a unique representation. And the representation of that object is unique in respect of its picking out the object in terms of the object's unique location.

Once a tracking mechanism is assigned to a target object in the above way, it could track that object in a way similar to the way Pylyshyn and Storm's simulation tracked objects. Each tracking mechanism could operate by the following commands:

- 1) While the targets are visually distinct, scan attention to a single target and encode its location. Then, when targets begin to move, do steps 2-4.
- 2) Scan attention to the location at which the target object was represented as being. Find the closest object to that location.
- 3) Update the representation of the target object with the location of the object found in (2). This becomes the new represented location of the target object.
- 4) Repeat from (2) until elements stop moving. Scan attention to the final encoded location and report the object there.

One's four or five tracking mechanisms could simultaneously track objects in this way. So, unlike in Pylyshyn and Storm's simulation, tracking would not involve a single tracking mechanism that shifts from one target object to the next. As a result, this representational model of tracking, like Pylyshyn's visual-index model of MOT, would avoid the tracking errors introduced by such shifting from object to object, as well as those errors introduced by the limited scanning speed

of a single, serial tracking mechanism. There would of course still be errors in tracking, as one would expect. For example, objects in MOT trials often cross paths. When this happens, it could be that an object other than the target object appears closest to where vision represented the target object a moment earlier. If so, the tracking mechanism would erroneously encode the location of that object as the new location of the target object, and it would subsequently begin tracking that new object. However, the tracking mechanisms could reduce such errors by using encoded information about their target objects' prior trajectories.

Since MOT could rest on such parallel mechanisms that track objects by correlating unique representations of those objects, Pylyshyn and Storm fail to show that MOT rests on a primitive, nonrepresentational access to objects. And, since Pylyshyn's view that vision binds representations of distinct properties of objects rests on the existence of such nonrepresentational access to objects, his view of feature binding is unmotivated.

Further, it is unclear how vision could access objects independently of representing their features. Pylyshyn claims that visual indexes are mental analogues of linguistic demonstratives, such as 'this' and 'that' (2003, p. 206). And linguistic demonstratives, he supposes, refer to objects independently of ascribing properties to them; they are supposed to be paradigms of nondescriptive, nonrepresentational reference. If so, they provide a model for how visual indexes could pick out target objects independently of representing any of their properties.

But Pylyshyn does not explain how demonstratives refer without representing their referents. And without a positive account of how they do so, demonstratives do not provide a useful model for visual indexes.

Further, it could be that demonstratives do in fact refer descriptively. Demonstratives could be disguised descriptions, just as Quine (1953) argues proper names and other singular terms are. According to Quine, singular terms, such as 'Henry Fonda', are to be regimented as definite descriptions, e.g., 'the lead actor in "Twelve Angry Men"'. Likewise, it could be that when one says, "That's nice," 'that' is to be regimented as some definite description, e.g., as 'the vase on the table'. On this view, demonstrative utterances express descriptive thoughts.

One reason to think demonstratives are in fact disguised descriptions is that one can always describe what one is referring to demonstratively. If one says, "That's nice," and one is asked what one is referring to, one can describe it, e.g., by saying, "The vase on the table is nice." This holds even for cases in which one is unclear about the nature of the thing one is referring to; e.g., when one is talking about some piece of technical equipment in a chemistry lab, one can at least describe it in terms of its location, shape, and color.

However, Pylyshyn cites an argument due to John Perry (1979) to argue that demonstratives do in fact refer independently of any description

or representation. If Perry's argument succeeds, then arguably demonstratives do refer nondescriptively, even if we have no account of how they do so.

The debate over the nature of demonstrative reference is extensive, and I will not attempt to settle it here. However, I will briefly discuss Perry's argument, and a reply on behalf of the descriptive theory of demonstratives. My aim here is to show that linguistic demonstratives do not clearly provide a useful model for Pyslyhyn's visual indexes.

In Perry's example, a hiker is looking for the Mt. Tallac trail, which leads out of the woods. The hiker, facing a trail, wonders whether it is the Mt. Tallac trail. Suddenly, the hiker begins to follow the trail, reflecting that the hiker has come to believe that the trail is in fact the Mt. Tallac trail.

According to Perry, "If asked, [the hiker] would have to explain the crucial change in his beliefs in this way: 'I came to believe that this is the Mt. Tallac trail ...'" (1979, p. 4; italics in original). The hiker, Perry claims, could not describe what he came to believe is the Mt. Tallac trail.

But the hiker could explain his change in belief descriptively, without any demonstratives. Perhaps when the hiker is wondering whether the trail is the Mt. Tallac trail, the hiker thinks about the trail as the trail straight ahead, i.e., the hiker thinks about the trail under the description 'the trail straight ahead'. The hiker's change in belief occurs when the hiker, who is looking for the Mt. Tallac trail, suddenly identifies

the trail straight ahead with the Mt. Tallac trail. This thought process involves no mental analogues of demonstratives. And when the hiker is asked to explain the change in belief, the hiker could describe it this way: “I came to believe that the trail straight ahead is the Mt. Tallac trail.” So Perry’s example does not show that demonstratives, or the mental analogues they express, refer independently of descriptions, or descriptive thoughts.

Without establishing that demonstratives refer nondescriptively, demonstratives fail to serve as a model for direct, nonrepresentational access. Since Pylyshyn does not explain how visual indexes, or demonstratives for that matter, pick out objects nonrepresentationally, and since nonrepresentational access to objects is not required for MOT, Pylyshyn’s argument that vision accesses objects independently of representing any of their properties fails. Since Pylyshyn’s object-based account of how we see feature conjunctions rests on such nonrepresentational access to objects, and since Pylyshyn has failed to establish that vision has such primitive access to objects, his account is unmotivated.

#### 10. Visual Indexes

But Pylyshyn does provide a positive account of how visual indexes pick out objects. On Pylyshyn’s view, a visual index picks out a target object in virtue

of a causal relation. A visual index is assigned to a target object when that target causes the activation of the visual index, e.g., when the target briefly flashes in the beginning of a MOT trial. And, since this relationship between the visual index and the target is causal, the visual index continues to pick out the target as long as the target continues to activate the visual index. Further, since the relation between the visual index and the object is not representational, it can be maintained while the object changes, e.g., while it moves.

Of course, this account of object tracking does not by itself explain how one sees an object's properties, e.g., when one sees a target as a small, blue circle. According to Pylyshyn, once a visual index is assigned to an object, vision sends detection signals back to it to determine its properties, e.g., color, shape, size, and location (2003, pp. 270-5).<sup>67</sup> Detecting these properties enables vision to construct a representation of the object as having the various properties detected. Pylyshyn claims that one sees feature conjunctions when vision forms representations of distinct properties in connection with the same visual index. Accordingly, one sees something as both red and square when vision forms a representation of red and a representation of square in connection with the same visual index, not when vision forms a representation of red and a representation of square at the same location, as Clark argues.

Further, since the visual index maintains its causal link with the object as the object moves, vision can continue to send detection signals to the object.

---

<sup>67</sup> More precisely, Pylyshyn claims vision sends detection signals to the proximal stimulus, not the distal stimulus.

This enables vision to update the representation of the object without having to correlate new and old representations, thus sidestepping the problem Pylyshyn raises for the view that vision accesses objects only in virtue of representing their properties.

But Pylyshyn's positive account of nonrepresentational visual access to objects fails for the same reason his analogy between demonstrative reference and visual indexing fails. A visual stimulus does not exist independent of its properties. Whenever a visual stimulus is in front of one's eyes, so are its properties. So it isn't clear how the stimulus, but not its properties, could causally activate the visual index. And if the visual index picks out whatever causally activates it, it isn't clear how it could pick out the object independent of its properties, as Pylyshyn argues. Further, since the target object does not exist independent of its properties, whenever vision sends a detection signal to the object, it sends it to its properties, e.g., it sends the detection signal to the location of the object. So it is not clear how vision could detect the properties of the object independent of detecting properties at the object's location.

There are two other problems one might also raise for Pylyshyn's account of visual indexes. I'll first examine a concern about the continuous causal connection that, according to Pylyshyn, holds between visual indexes and objects. One can track objects through brief disruptions in the causal connection between those objects and vision, i.e., when one blinks, or when objects are briefly occluded by other objects. Since Pylyshyn claims vision tracks objects by

maintaining a continual causal connection with them, such disruptions in the causal connection could pose a problem for his view. However, I'll show that Pylyshyn's view withstands this criticism. I'll then raise a problem for Pylyshyn's account of how vision forms representations of the properties of the objects it is tracking. One might argue that vision could not send detection signals to an object without first representing the object's location, as Pylyshyn claims. But if Pylyshyn's account of binding rests on visual representations of objects' locations, it is a location-based account. I'll argue that Pylyshyn can circumvent this objection only if he can show that we never see feature conjunctions of properties of objects to which vision has not assigned visual indexes. But I'll discuss experimental data suggesting that vision not only represents features of objects it is not tracking but binds those representations. So, even if they are required for tracking multiple objects, visual indexes are not required for one to see conjunctions of features.

#### 11. The Problem of Tracking Despite Causal Interruptions

Pylyshyn argues that we track objects in virtue of their continuously causing the activation of visual indexes. It is because this connection is purely causal and requires no representations of the object one is tracking that one can track an object while it moves. And it is because the connection is continuous that one can track the object for an extended period of time.

But we can track objects despite interruptions in the causal connection between them and vision. For example, one can continue tracking an object even after it is briefly occluded by another object, or after one blinks or saccades. So Pylyshyn must explain how one tracks objects across such interruptions in the causal connection.

Perhaps Pylyshyn can do so with only a minor modification of his view. Causal mechanisms often continue functioning after interruptions in causal connections. For example, trains move in virtue of electrical charges running through tracks or wires to which the trains are connected. But if the power running through the track or wire to the train is momentarily interrupted, e.g., when the train runs over a length of dead track, the train continues to move, since it has momentum. As long as the train has enough momentum to roll past the dead track, the power will again cause the train to move once the train reaches a live stretch of track. Likewise, visual indexes could have properties analogous to momentum in virtue of which they stay activated, or resonate, for a short period after a target disappears. As long as the target then reappears close enough to where it disappeared, and before the visual index stops resonating, that target will resume causing the activation of that visual index.

This account suggests that a visual index will reattach itself to whatever object is close enough to where the target disappeared. But in some cases two objects are both close enough to where the target was when it disappeared. In such cases, both objects are equally good candidates for tracking on this view. If

one successfully continues to track the right object in these cases, we must explain how one does so.

One might argue that the visual index attaches to whichever object is both close enough and closest to where the target was just before the interruption. But this would have surprising consequences. Suppose one is tracking a target that disappears behind an occluder. While the target is occluded, another object moves towards the location where the target disappeared. When the target reappears, the other object is located where the target was just before it disappeared, and the target is on the other side of the occluder (figure 7). If one successfully continues tracking the target in such cases, visual indexes do not simply reattach to whatever object is closest to where the target was just before it disappeared, since the distractor, not the target, is the object closest to that location.

Figure 7

	L1	L2	L3	L4	
T1:	D→		X→	□	(target about to disappear)
T2:		D→		□	(target behind occluder)
T3:			D→	□→X	(target reappearing from occluder)

Note: 'X' represents the target. 'D' represents the distractor, an object one isn't tracking that has all the same properties, other than location, as the target.

At time T1, the distractor is at location L1, and the target is at L3, about to go behind the occluder at L4. At T2, the distractor is at L2 and the target is behind the occluder at L4. At T3, the distractor is at L3, where the target was just before being occluded, and the target is on the other side of the occluder.

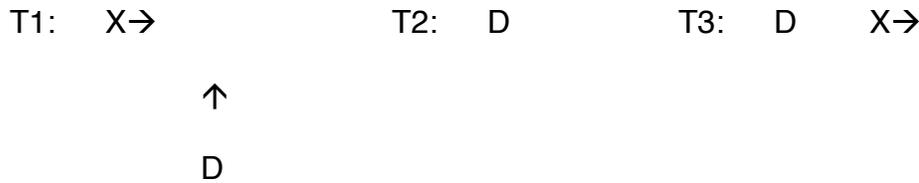
Also, imagine a case in which, when one blinks or saccades, another object moves to where the target was and the target continues along its trajectory. Presumably, one continues to track the target even though the other object is now closest to where the target was when it disappeared (figure 8). If one continues tracking the target in such cases, we must explain how one does this, since the distractor, not the target, is closest to where the target was when it disappeared, i.e., when one blinked or saccaded.

Figure 8

	L1	L2	L3
T1:	D→	X→	
T2:		(blink)	
T3:		D→	X→

At T1, the distractor is at L1 and the target is at L2. At T2, the subject blinks or saccades. At T3, when one opens one's eyes or stops saccading, the distractor is at L2, where the target was before the blink or saccade, and the target is at L3.

Finally, suppose a case in which two objects moving at different speeds converge on the same point, one occludes the other, and they then continue moving in the directions in which they were moving before the occlusion (figure 9). If one is tracking the faster of these two objects, then from the moment they begin to separate, the distractor will be closest to the location at which the two objects met. If one continues to track the target, one must do so with respect to something other than its proximity to the point at which it met the other object.

Figure 9

At T1, the distractor is moving up towards the location of the target, and the target is moving towards the right. At T2, the distractor occludes the target. At T3, the distractor, which is slower than the target, is closer to where they overlapped than the target is.

But it could be that the visual index reattaches to the target in such cases in a way analogous to the way the train reattaches to its power source after running over a stretch of dead track. If the visual index has a property analogous to motion (VI-motion), and the objects in these special cases do not radically change direction, the visual index can reattach to the target, just as the train reattaches to its power source after rolling past the dead track. In the first example (figure 7), while the target is occluded the visual index continues VI-moving in a way analogous to the way the target was moving before it was occluded. In the second example (figure 8), the visual index VI-moves in the same way during the blink. And in the third example (figure 9), the visual index VI-moves in a way analogous to the way the target was moving before being occluded by the distractor.

But proximity to the location at which a target disappeared, direction of movement, and velocity are not the only spatiotemporal features that affect one's

continuing to track an object after an interruption in the causal connection between the target and vision. Brian Scholl and Pylyshyn (1999) showed that tracking is also sensitive to the way targets disappear when being occluded by other objects. Tracking is not significantly impaired when a target is occluded, as long as the target deletes and accretes along a fixed leading contour, i.e., when the target appears to gradually disappear behind, and then reappear from behind, the occluder. But when the target disappears or reappears instantaneously, or it deletes and accretes gradually along the wrong fixed contour, continued tracking is significantly impaired. This suggests tracking is sensitive to certain pictorial depth cues that indicate an object's disappearance behind, and then reappearance from behind, another object. Such sensitivity shows that properties of the visual index analogous to motion are not always sufficient to maintain tracking through disruptions in causal connections between targets and visual indexes.

Lavanya Viswanathan and Ennio Mingolla (2002) also showed that, though subjects can continue tracking a target after it is occluded by a distractor, they are significantly better at doing so when they are provided with depth cues indicating which of the objects, the target or the distractor, is being occluded. In such cases, when the target begins to reappear from behind the distractor, both objects are equally close to the location at which the target disappeared. So vision must differentiate between the target and the distractor based on something other than proximity to that location. In part, properties of the visual

index analogous to motion could help vision differentiate between the target and the distractor, especially if the target continues along the same trajectory it was following before it was occluded. But we cannot explain the effects of the depth cues along such lines.

Perhaps continuing to track an object after it is occluded by another object depends on representations of the boundaries of the two objects. On this view, vision determines which, the target or the distractor, is being occluded on the basis of the objects' visible boundaries. When vision represents the target's visible boundaries as changing in a certain way, whereas it represents no change in the visible boundaries of the distractor, it determines that the distractor is occluding the target. When vision represents certain kinds of changes in the distractor's visible boundaries, while representing no changes in the target's visible boundaries, it determines that the target is occluding the distractor. If so, continuing to track a target after a brief disruption in the causal connection between the target and vision requires representations of objects' properties in at least some cases.

On the other hand, perhaps we can explain the sensitivity of visual indexes to pictorial depth cues in a way similar to the way we explain their sensitivity to an object's location. We could explain the effect of pictorial depth cues on tracking in terms of subpersonal, nonrepresentational properties of the visual index. To explain the sensitivity of visual indexes to pictorial depth cues we need not invoke representations of objects' visible boundaries any more than

we need to invoke representations of the shapes, sizes, and weights of pieces of debris to explain a vacuum cleaner's sensitivity to the shapes, sizes, and weights of pieces of debris. So neither Scholl and Pylyshyn's nor Viswanathan and Mingolla's results pose a problem for Pylyshyn's nonrepresentational account of object tracking.

## 12. The Problem of Detecting Features

According to Pylyshyn, once a visual index is assigned to the object that causes its activation, and thus picks out that object, vision sends detection signals back to that object to determine what properties it has. Once vision detects the object's properties, it constructs a representation of the object as having those properties. On this view, binding occurs when representations of distinct features are formed in connection with the same visual index; those representations then represent those distinct features as features of the same object, so one sees a feature conjunction.

But if vision must send detection signals to an object for one to see feature conjunctions, Pylyshyn owes an account of how vision does this without first representing the object's location. Without such an account, his view fails to provide an alternative to Clark's and Treisman's location-based views.

According to Pylyshyn, vision can send a detection signal to an object without representing the object's location if it first detects properties correlated with the object's location. To show that vision can do this, Pylyshyn cites the

case of a baseball player running to catch a fly ball. Rather than computing and representing where and when the ball will land, the player detects correlates of the ball's destination. The player "... moves so as to nullify the apparent curvature of the ball's flight, so it looks like it is descending in a continuous straight line (McBeath, Shaffer, and Kaiser, 1995)" (2003, p. 221).<sup>68</sup> By moving so as to manipulate properties correlated with the ball's destination, the fielder is able to run to the right place to catch the ball. He needs no representation of that location to do so, e.g., no representation of the ball as destined for a place on the foul line, 20 ft. from the left field fence, and 198 ft. from third base.

But, one might argue, if the player gets to the ball by detecting properties correlated with its destination, we must explain how he detects those correlates. If vision detects the correlates of location by sending detection signals to the ball, then we must explain how vision does that without first detecting the ball's location or other correlates of its location. If detecting a feature such as the apparent curvature of the ball's trajectory requires prior detection of other correlates of the ball's location, which are required to direct the detection signal, there will be a regress of feature detection. And if vision sends the detection

---

<sup>68</sup> Peter McLeod, Nick Reed, and Zoltan Dienes (2002) argue that fielders do this in virtue of detecting the ball's apparent vertical acceleration, not the apparent curvature of its trajectory.

signal to the ball in virtue of first detecting the ball's location, Pylyshyn's view of binding is location based.<sup>69</sup>

Though Pylyshyn's example addresses the issue of how one performs an action, i.e., running to a particular place to catch a ball, it is also relevant to the issue of how vision sends a detection signal to an object to enable one to see it. According to Pylyshyn, the baseball player can move to where the ball will land without having a visual representation of that place. He does this by detecting features of the ball that correlate with its destination. Likewise, perhaps vision could detect correlates of a target object's location and use those correlates to direct a detection signal to the object.

However, just as the baseball player must already detect some feature of the ball to detect the correlates of the ball's location, vision must already detect some correlate of an object's location to detect the correlates of location it uses to direct the signal used to perform binding. So the regress problem applies equally to the case of vision's sending a detection signal to an object as it does to the case of one's running to catch a foul ball.

Perhaps Pylyshyn could avoid the regress by claiming that subpersonal states, e.g., those enabling figure-ground segmentation, encode information about an object's location, that those subpersonal states are used to direct detection signals, and that such subpersonal states do occur prior to the

---

<sup>69</sup> Of course, Pylyshyn's location-based view could differ in important respects from Clark's and Treisman's view.

assignment of a visual index to an object.<sup>70</sup> Accordingly, vision could track objects and bind representations of their various properties without forming personal-level representations of the locations of those objects. If so, perhaps MOT and binding require no personal-level sensory representations of objects' locations.

Such subpersonal states parse the visual scene by grouping visual features into units roughly corresponding to distinct visual stimuli. Elements clustered together in one's field of view are often grouped in this way, e.g., when dots positioned close together appear to form a row or column. Presumably, vision must encode or register at least an object's location, shape, and size to distinguish it from the background and other objects. And vision can encode these features without sending detection signals to objects; grouping mechanisms in early vision could operate on information from incoming signals. If so, the subpersonal states involved in this grouping operation could make location information available to direct and send detection signals used to construct a personal-level representation of an object, and to update that representation as the object changes.<sup>71</sup> Pylyshyn could then claim that one sees feature conjunctions only once vision forms personal-level representations of objects, and that occurs only after vision assigns visual indexes.

---

<sup>70</sup> This is amenable to Pylyshyn's claim that such figure-ground segmentation is a prerequisite for all visual operations (unpublished, p. 18).

<sup>71</sup> Pylyshyn, himself, suggests a similar solution (2003, p. 273).

But if vision can send a detection signal to an object, and thus perform binding, only after subpersonally registering the object's location, then binding rests on that registration of location. If so, one sees a conjunction of two features, e.g., color and shape, partly in virtue of detecting them at the same location. Therefore, this response would commit Pylyshyn to a location-based view of binding, albeit a subpersonal version.

Nevertheless, since such subpersonal encodings of location are not personal-level visual representations, or sensations, of objects' locations, such subpersonal states might enable vision to send detection signals without first enabling one to see an object as being, e.g., off to the left. If so, Pylyshyn could argue, the involvement of such subpersonal encodings of location does not undermine his object-based view.

So it could be that binding occurs when a visual index is assigned to an object, and the object's features, including its location, are then encoded in a bound, personal-level representation. In this case, personal-level representations of location play no special role in binding. So, perhaps, one sees feature conjunctions when vision forms personal-level representations that represent distinct features as features of the same object.

But one might argue that this subpersonal solution to the feature-detection problem faces the same problem Pylyshyn raises against the representational account of object tracking. Since the objects one is tracking in MOT tasks are moving, vision would use different subpersonal states to direct feature-detection

signals to the same objects at different moments, i.e., to update a representation of a changing, moving object. But to do this, vision must determine which subpersonal states register the different locations of the same object at different times. At time T1, vision can send a detection signal using subpersonal state S1, which carries information about location L1, where a target object is located. Vision then compiles a representation of the object as having the various features detected. But by time T2, the object has moved from L1 to L2. So to send detection signals and update the representation it formed at T1, vision must use a different subpersonal state, S2, encoding L2. To use subpersonal state S2, vision must determine that S2 registers location information about the target. But S2 carries different information from S1, the subpersonal state vision used to send a detection signal to the target at T1, and it carries different information from that encoded in the personal-level representation of the target object. Since there are subpersonal states registering the locations of distractors as well as targets, vision must somehow determine which subpersonal state picks out the location of the target object to send a detection signal to an object it is tracking.

Vision could select the right subpersonal state by selecting the one carrying information about the location both closest and close enough to the location to which it sent a detection signal a moment before.<sup>72</sup> To do this, vision must detect the relative distances between locations registered by subpersonal

---

<sup>72</sup> Pylyshyn suggests that there could be a mechanism that performs such a function (2003, pp. 273-274).

states at successive moments, and it must store information about where it sent the detection signals a moment earlier.

But this is just a subpersonal version of the proposal I introduced to explain MOT in terms of personal-level representations of objects' locations. If Pylyshyn invokes such an account to defend his view, he must explain why it could hold at the subpersonal level but not at the personal level.

Alternatively, instead of invoking subpersonal states that encode objects' locations, perhaps Pylyshyn could avoid the regress of detecting correlates of an object's location by abandoning the view that vision assigns a visual index to an object before detecting that object's properties. Rather, vision could assign a visual index to an object and encode that object's properties at the same time. If so, there is no need to send a detection signal to perform binding, so no regress of detection signals.

But if visual indexes are not assigned to objects before vision constructs representations of those objects, and thus performs binding, Pylyshyn must explain what visual indexes have to do with binding. Perhaps the representations of an object's properties depend no more on the visual index than the visual index depends on the representations.

Perhaps Pylyshyn could argue that updating the representation as the object changes does rely on the visual index. To update a representation of an object as the object changes, he might claim, vision must continue to access the object, and visual indexes provide such continuous access.

But visual indexes enable continuous access to objects in virtue of being continuously caused by them. And if an object causes visual representations of itself at the same time it activates visual indexes, we can explain how vision updates a representation of an object in terms of the object's continuously causing new representations of itself. If the object changes, it causes a representation different from the one it caused a moment ago. Again, visual indexes might have nothing to do with it.

However, if the only feature conjunctions one sees are conjunctions of properties of objects to which vision has attached visual indexes, then perhaps visual indexes are in fact required for one to see feature conjunctions. Pylyshyn's argument for object-based binding thus rests on the assumption that the only representations that get bound are those that represent properties of objects to which vision assigns visual indexes.

If one sees conjunctions of properties of objects vision isn't tracking, or to which it has not assigned a visual index, then even if visual indexes are required for tracking, they have nothing to do with one's seeing feature conjunctions.

### 13. Vision Encodes Properties of Objects It Isn't Tracking

Recent experiments on implicit visual processing during MOT show that vision encodes features of objects it isn't tracking and that one sees not only individual features of those objects but conjunctions of those features. Since, according to Pylyshyn's view, visual indexes are

not assigned to the objects one is not tracking, these recent experiments show that visual indexes are not required for one to see feature conjunctions. Hirokazu Ogawa and Akihiro Yagi (2002) showed that one's performance in MOT tasks improves more over the course of a series of trials when the movements of the distractors are the same across those trials than when the movements of the distractors are different from trial to trial. This suggests that subjects do in fact see and remember the movements of the distractors, even though they are not tracking those distractors. Since, according to Pylyshyn, vision does not assign visual indexes to these distractors, visual indexes are not required for representing objects' movements. So seeing an object's movements does not require the primitive access to objects Pylyshyn supposes visual indexes provide.

Ogawa and Yagi ran three different MOT experiments. In each of these experiments, subjects performed five consecutive MOT trials, tracking five of ten identical objects in each trial. In the first experiment, the all-new phase, the movements of all objects differed from trial to trial. In the second, the old-target phase, the movements of the targets were invariant from trial to trial, but the movements of the distractors varied from trial to trial; i.e., the targets each moved in exactly the same ways as in the previous trials, but the distractors did not. And in the third, the all-old phase, the movements of all objects, both targets and distractors, were invariant from trial to trial.

Subjects' tracking performance improved more in the old-target trials than in the all-new trials, and their tracking performance improved more in the all-old trials than in the old-target trials. This shows it is easier for subjects to track targets when distractors' movements are invariant from trial to trial. This, in turn, shows that subjects both see and remember the movements of distractors.

However, when asked whether there were invariances in movement patterns across trials, subjects' reports were at chance; they were not aware that they saw the same patterns of movement across trials. So, though subjects see and remember the distractors' movements, they do not see or remember those movements consciously.

These experiments show vision represents and stores representations of distractors' movements. Since, according to Pylyshyn, vision does not assign visual indexes to distractors, this shows that assigning visual indexes is not required for representing at least some of an object's properties. If vision represents objects' properties independent of visual indexes, perhaps it binds representations of them too.

However, it could be that some representations are formed in connection with visual indexes and only those representations get bound. So, perhaps, binding depends on visual indexes, even if vision forms representations of objects' features independent of them.

But recent experiments conducted by Brian Scholl, Pylyshyn, and Steven Franconeri (unpublished) suggest that vision not only forms

representations of distractors' features but also binds those representations. These experiments thus show that, even if visual indexes are required for MOT, they are not required for binding. So, even if vision uses nonrepresentational access to objects to track them, it doesn't use that same access to bind representations of their features. If so, Pylyshyn fails to establish that binding is object based, not location based.

Scholl et al. tested whether subjects could recall the colors and shapes of both targets and distractors. Subjects tracked four of eight objects. Each object could be one of three shapes and one of three colors; i.e., an object could be red, yellow, or blue and it could be T-shaped, +-shaped, or L-shaped. Further, at various times during the trials, objects could disappear behind occluders. During those periods, there was a 50% chance that one of the objects would change either its color or shape.

At the end of each trial, all of the objects disappeared; then all but one reappeared. In place of the missing object was a marker, and subjects were to report what color and shape the missing object was before it disappeared.

Scholl et al. found subjects were no better at recalling the colors and shapes of targets than those of distractors. However, subjects' reports of objects' colors and shapes were above chance for both targets

and distractors, suggesting that subjects did in fact see the colors and shapes of both targets and distractors.

This result, like Ogawa and Yagi's, shows that vision does in fact represent properties of objects it is not tracking. Again, since subjects have not assigned visual indexes to objects they are not tracking, even if tracking requires visual indexes, representing properties does not.

Further, that subjects reported both the colors and shapes of distractors above chance shows that they saw conjunctions of distractors' properties. This in turn shows that vision binds representations of features of objects it is not tracking. So, even if tracking requires visual indexes, binding does not. Contrary to Pylyshyn's argument, the results of MOT experiments do not support the view that binding is object based, not location based.

But Scholl et al. argue it could be that subjects report the colors and shapes of distractors above chance because they allocate some attention to distractors in addition to targets (unpublished, p. 14). Tracking four of eight objects is easy for some subjects. So perhaps they track more than just the designated four targets. Further, objects sometimes become concentrated in one region of the screen, enabling subjects to momentarily allocate attention to distractors. According to Scholl et al., both of these factors could explain how vision is able to encode the colors and shapes of some distractors. If so, the above-chance reports of distractors' properties do not show that vision encodes

properties, nor that it binds representations of distinct properties, independently of assigning visual indexes to objects.

But subjects report the prior colors and shapes of distractors as successfully as those of targets. So, if subjects' above-chance reports result from their attending to and tracking distractors, they must attend to and track distractors and targets equally. But subjects do not attend to every distractor for the entire duration of the trial, and they do not track all of the distractors. This is confirmed by the results of two other experiments in which Scholl et al. found that subjects recall the previous locations and directions of movement of targets significantly better than those of distractors. Presumably, if subjects tracked and attended to distractors and targets equally, they would recall distractors' and targets' prior locations and directions of movement equally well. On the other hand, if subjects do not track and attend to distractors and targets equally, as these experiments suggest, we cannot account for why they recall the colors and shapes of distractors equally well as those of targets in terms of their attending to and tracking distractors.

Since subjects do in fact see conjunctions of features of objects they are not tracking, binding does not require visual indexes, even if MOT does. So Pylyshyn's object-based account of binding is unmotivated and fails to undermine Clark's location-based view.

#### 14. Problems With Clark's Location-Based Binding

However, this does not show that Clark's theory of sensing feature conjunctions is right. In this section, I'll argue that Clark's view rests on the false assumption that mental qualities occur separately and therefore need binding. In so arguing, I'll offer an alternative account of sensing feature conjunctions, according to which distinct mental qualities are interdependent and therefore need no binding. On this view, which is a consequence of homomorphism theory, sensing the locations of properties plays no special role in sensing feature conjunctions.

Clark's view that we sense feature conjunctions by sensing distinct features at the same location rests on the assumption that mental qualities occur independently of each other. If mental qualities, such as those in virtue of which one sees color and shape, do not occur independently of each other, then they need not be bound for one to see a combination of color and shape. And if distinct mental qualities need not be bound, then we need not commit to a special mechanism of sensory localization, such as a sensory name, to explain how distinct mental qualities are bound.<sup>73</sup> So, if distinct mental qualities do not occur independently of each other, it could be that one senses an object's location in virtue of having a sensation with a mental quality corresponding to that location, i.e., a location\*, as homomorphism theory holds.

---

<sup>73</sup> Likewise, we need not commit to the involvement of a mechanism, such as a visual index, that picks out an object independent of its properties, as Pylyshyn argues.

Clark assumes mental qualities corresponding to different kinds of perceptible properties, such as color, shape, and size, occur independently of each other, and therefore need binding, because we can sense different combinations of perceptible properties; e.g., we can see red squares, red triangles, green squares, and green triangles, all of various sizes and orientations. Clark raises the many-properties problem to illustrate this point. Since we can see different combinations of perceptible properties, Clark assumes there is some mechanism in virtue of which different mental qualities are combined.

But there is good reason to think mental qualities are interdependent in such a way that they do not require binding, even though we can sense different combinations of perceptible properties. One never sees an object's color without also seeing its shape and size. All colored surfaces appear to be spatially extended and to have boundaries. And one sees such surfaces in virtue of having sensations with colors\*, shapes\*, and sizes\*. So one never has colored\* sensations that have no shape\* or size\*. Likewise, one never sees an object's shape without seeing some color or size. So one never has a visual sensation with a shape\* but no color\* or size\*.

We can explain the interdependence of distinct mental qualities in terms of homomorphism theory, the view that mental qualities are mental analogues of perceptible properties. Because they are analogues of perceptible properties, mental qualities bear many of the same relations to each other that distinct

perceptible properties bear to each other. So colors\*, visual shapes\*, and visual sizes\* could relate to each other in ways parallel to the ways colors, visible shapes, and visible sizes relate to each other.<sup>74</sup> For example, there are no visible colored surfaces that have no shape or size because the visible shape and size of a surface are determined by the visible boundaries of the color of the surface. And since all colored surfaces are spatially extended and have boundaries, all colored surfaces have some shape and size.<sup>75</sup> So color, visible shape, and visible size are all interdependent.

Colors\*, visual shapes\*, and visual sizes\* could relate to each other in an analogous way. Colors\* are mental analogues of colors, so they bear the same relations to shapes\* and sizes\* that colors bear to visible shapes and visible sizes. So colors\*, visual shapes\*, and visual sizes\* are interdependent, just as colors, visible shapes, and visible sizes are interdependent. Specifically, just as the boundaries of colors determine the visible shapes and sizes of colored surfaces, the boundaries of colors\* determine the visual shapes\* and sizes\* of visual sensations (Rosenthal, 2005). According to this view, we sense feature

---

<sup>74</sup> I specify that these shapes\* and sizes\* are visual shapes\* and visual sizes\* because tactile sensations also have shapes\* and sizes\*, and the shapes\* and sizes\* of visual and tactile sensations could be distinct. In the next chapter, I argue that the mental qualities in virtue of which we sense objects' spatial properties are in fact distinct in different modalities.

<sup>75</sup> One might argue that this view fails to account for all cases of seeing color, e.g., when one looks up at a clear, blue sky, or when one sees a Ganzfeld. These cases, one might argue, are cases of unbound color, so color without shape or size. But in these cases, the color one sees is bounded by the limits of one's field of view, the space beyond which one sees nothing at a given moment. So the colored expanse one sees is the shape of one's field of view.

conjunctions because distinct mental qualities are interdependent in a way that parallels the way distinct perceptible properties are interdependent.

This view solves the so-called many-properties problem. When one sees a red square next to a green triangle, one has a red\* sensation the boundary of which determines the mental quality square\* and a green\* sensation the boundary of which determines the mental quality triangular\*. When one sees a green square next to a red triangle, the boundaries of one's green\* sensation determine the mental quality square\*, and the boundaries of one's red\* sensation determine the mental quality triangular\*. So we need not hold that distinct mental qualities, such as color\* and shape\*, are independent of each other in order to solve the many-properties problem, as Clark assumes.

Since we can explain how we sense feature conjunctions in terms of the interdependence of distinct mental qualities, Clark's location-based view is superfluous, so we do not need a special account of how we sense objects' locations, as Clark argues. Rather, we sense objects' locations in virtue of having sensations with mental qualities that represent the sensible locations of those objects. And those mental qualities, locations\*, represent perceptible locations of stimuli in virtue of resembling and differing from each other in ways parallel to the ways those perceptible locations resemble and differ from each other. For example, just as two objects off to the left in one's field of view are more similar with respect to their horizontal position in one's field of view than either is to an object off to the right, one's sensations of those two objects are

more similar to each other than either is to the sensation of the object off to the right. A sensation of something off to the left in one's field of view is visually off-to-the-left\*, and a sensation of something off to the right in one's field of view is visually off-to-the-right\*.

Further, the mental qualities representing sensible locations are themselves inseparable from other mental qualities, such as colors\*. One never sees a location without seeing something with a color, shape, and size at that location, nor does one ever see the color, shape, or size of something without seeing it at some location.

Clark might argue that neuroscience shows distinct mental qualities are in fact separate and need binding for one to sense feature conjunctions. Visual processing from primary visual cortex, V1, projects forward to higher areas of visual cortex, e.g., V2-MT/V5. And those higher cortical areas process distinct visible properties, e.g., color, shape, orientation, and motion, separately. Clark identifies mental qualities with the features of those separate neural representations (2000, p. 44). Accordingly, some feature of the neural representation one has when seeing a red square is identical with the mental quality red\*, and that neural representation occurs in a different area of visual cortex from that of the neural representation with the feature identical with the mental quality square\*. Since the neural representations of distinct perceptible properties occur in separate areas of visual cortex, Clark thinks the mental qualities are separate. If so, just as the neural representations of distinct visible

properties must be bound to enable one to sense a feature conjunction, e.g., of color and shape, the mental qualities representing those distinct properties must be bound.

But it is unclear why Clark identifies features of the separate neural representations in areas V2-MT/V5 with mental qualities. Mental qualities, such as red\* and square\*, are folk-psychological posits, posited to explain the qualitative characters of sensations. Red\*, e.g., is the mental quality in virtue of which a sensation of red is that particular kind of sensation and not a sensation of some other color. And, again, one never sees color, e.g., without also seeing shape and size. So one never has sensations of color without sensations of shape or size. This suggests that one never has a sensation with color\* but no shape\* or size\*. Likewise, one never sees shape without seeing color or size, so one never has a visual sensation with shape\* but no color\* or size\*. Arguably, one never has a sensation with a single, solitary mental quality; rather, mental qualities are interdependent and do not require binding, just as perceptible colors and perceptible shapes do not require binding. If neural representations of distinct perceptible properties are separate, i.e., in virtue of occurring in separate areas of visual cortex, then those neural representations are not identical with sensations, and the features of those neural representations are not identical with mental qualities. In fact, the binding problem in neuroscience is particularly interesting, not only because separate neural representations of distinct perceptible properties somehow give rise to unified sensations of feature

conjunctions, but because the sensations they give rise to are always sensations of feature conjunctions.

Further, distinct perceptible properties, such as color, shape, size, orientation, and motion, are all represented together in both V1 and the lateral geniculate nucleus (LGN), a subcortical visual area that receives information directly from the optic nerve and projects to V1. So it could be that features of neural representations in the LGN or V1 are identical with mental qualities. If so, there is no reason to think the separate neural representations in other visual areas show that distinct mental qualities are separate and need binding.

Of course, those separate neural representations in areas V2-MT/V5 could be involved in sensing, even if they are not identical with sensations. It could be that features of neural representations in V1 are identical with mental qualities and the separate neural representations of distinct perceptible properties in areas V2-MT/V5 are subpersonal processing states that are necessary but not sufficient for sensing. If one sees something only once those separate neural representations are bound, there's no reason to identify any individual, unbound neural representation with a sensation, or any of the features of any of those neural representations with mental qualities, even if the separate neural representations are involved in sensing. So there's no reason to claim mental qualities need binding just because the neural representations do.

In fact, in addition to the feedforward projections from V1 to other areas of visual cortex, there are feedback projections from those areas to V1 (Lamme,

2004; Bullier, 2001; Pascual-Leone & Walsh, 2001; Hupé et al. 1998; Cowey & Walsh, 2000).<sup>76</sup> This further suggests that the neural representations in visual areas V2-MT/V5 do not by themselves underlie visual sensations, but serve some intermediary role in sensing. Those higher visual areas could process information about color, shape, size, orientation, and motion to enhance representations underlying sensations in V1. Perhaps representations in V2-MT/V5 fine-tune processing of color, shape, orientation, size, and motion for more accurate sensing, or to enhance the segmentation of the visual scene into figure and ground (Hupé et al. 1998). Or, perhaps, those neural representations make one's visual sensations conscious (Bullier, 2001; Lamme, 2004; Pascual-Leone & Walsh, 2001). Whatever the roles of those separate neural representations of distinct visible properties, feedback from higher visual areas to V1 suggests that those separate neural representations are not themselves identical with sensations of the visible properties they process. If those neural representations are not identical with sensations, the features of those neural representations are not identical with mental qualities, as Clark assumes. So separate neural representations of color and shape, e.g., do not show that colors\* and shapes\* are separate and require binding.

In this chapter, I've focused primarily on visual cases of sensing feature conjunctions. But we sense feature conjunctions in all sensory modalities. For example, when one touches a doorknob, one feels something round, hard,

---

<sup>76</sup> There is also feedback from visual cortex to the LGN (Levine, 2000).

smooth, and cool; when one tastes a curry dish, one tastes something both sweet and spicy; and when one stubs one's toe, one has a sensation that's both painful and throbbing. Homomorphism theory explains all such cases of sensing feature conjunctions, not just the visual cases. One feels the doorknob as round, hard, smooth, and cool, e.g., in virtue of having a round\*, hard\*, smooth\*, and cool\* sensation, where those mental qualities are interdependent, just as the shape, solidity, texture, and temperature of the doorknob are interdependent. And though all of these mental qualities are also inseparable from the location\* in virtue of which one feels where the doorknob is located, it is not in virtue of their all having the same location\* that they are had in conjunction with each other.

## Chapter 5: The Qualitative Character of Spatial Perception Across Modalities

### 1. Introduction

We sense the spatial properties of objects in different sensory modalities. One can both see and feel the shape and size of an object, and one can see, feel, and hear where something is located. Likewise, one can feel various spatial properties of one's own body and of bodily stimulation, e.g., when feeling the movements and relative positions of one's own limbs, and when feeling tickles, itches, and pains. Since one senses the properties of objects in virtue of having states, e.g., sensations and perceptions, with mental qualities, one's visual, tactile, auditory, and bodily sensations all have mental qualities pertaining to spatial properties. To fully explain sensing, we must determine whether the mental qualities of sensations pertaining to the same spatial properties are themselves the same in different sensory modalities. We must determine whether, e.g., visual and tactile sensations of the same shape have some amodal property in common, in virtue of which both are sensations of that same shape.<sup>77</sup>

---

<sup>77</sup> Visual and tactile sensations of the same shape do of course have properties in common. For example, a visual sensation of a square and a tactile sensation of a square both have the property of representing an object as square. But it is not simply in virtue of representing a square that a sensation is a sensation of a square and not a sensation of some other shape. Thoughts about squares also represent squares, but they are not sensations of squares. Rather, a sensation of a square is a sensation of that shape in virtue of having a

In section 2, I discuss the relation between a solution to this problem and the account of sensing combinations of distinct properties I discussed in the previous chapter. I argue that it is a consequence of the view I argued for there that the mental qualities of visual and tactile sensations pertaining to the same shapes are distinct, modality-specific mental qualities. In section 3, I discuss and argue against John Campbell's (1996a, b) claim that the properties of visual and tactile sensations pertaining to shape are amodal. In sections 4 and 5, I further argue that the mental qualities of sensations pertaining to shape are different, and I offer homomorphism theory as the best account of the modality specificity of such mental qualities. In sections 6, 7, and 8, I discuss experiments from developmental psychology studying the ability of infants to recognize in one sensory modality, e.g., sight, spatial properties of objects they have previously sensed in another sensory modality, e.g., touch. Though one might argue that the results of these experiments show that such crossmodal shape recognition is innate, and that such innate abilities are best explained in terms of the amodality of the mental qualities pertaining to shape, I argue that these experiments fail to support either claim. In section 9, I examine further experiments on infants' ability to perform crossmodal shape recognition, or crossmodal shape transfer, that reveal certain asymmetries in these abilities at different stages of infant development. These results, I argue, support the view that the mental qualities

---

particular mental quality. The opposing views I am discussing are thus the view that visual and tactile sensations of the same shapes have amodal mental qualities that represent those shapes, and the view that they have modality-specific mental qualities that represent those shapes.

of visual and tactile sensations pertaining to the same shapes are in fact distinct. Finally, in section 10, I examine recent neurophysiological data cited as evidence for common, bimodal representations of shape involved in both tactile and visual shape perception. I argue this data does not show that the properties of visual and tactile sensations pertaining to shape are the same.

## 2. Feature Conjunctions and Modality Specificity

In the preceding chapter, I argued that the mental qualities of visual sensations pertaining to the shapes of objects, i.e., shapes\*, are determined by the mental analogues of boundaries of the mental qualities that represent the colors of objects, i.e., colors\*. If so, since tactile sensations of shape do not have colors\*, they do not have the same shapes\* as visual sensations.

Of course, tactile sensations have mental qualities pertaining to shape, since we do in fact feel shapes. And we can account for the shapes\* of tactile sensations in a way analogous to the way we explain the shapes\* of visual sensations. Just as the boundaries of colors\* determine the shapes\* of visual sensations, the boundaries of mental qualities pertaining to the textures, temperatures, and pressures of tactile stimuli, i.e., textures\*, temperatures\*, and pressures\*, determine the mental shapes of tactile sensations. On this view, visual and tactile sensations of the same shapes have different shapes\*, since

visual and tactile shapes\* are determined by different, modality-specific mental qualities.<sup>78</sup>

The view that shapes\* are modality specific is compatible with the homomorphism theory of sensing (Rosenthal, 1991, 2005; Meehan, 2002, 2003). According to homomorphism theory, mental qualities resemble and differ from each other in ways that parallel the ways the perceptible properties they represent resemble and differ from each other. Mental qualities thus represent their perceptible counterparts in virtue of homomorphisms between families of mental qualities and families of perceptible properties. And two distinct families of mental qualities could both be homomorphic to the same family of perceptible properties. So the family of visual shapes\* and the family of tactile shapes\* could be distinct, even though both are homomorphic to the same family of physical, perceptible shapes. And, since shapes\* are individuated in respect of their positions in their quality families, visual and tactile shapes\* could be distinct.

Of course, it could be that visual and tactile shapes\* are distinct even if the families of visual and tactile shapes\* are homomorphic to the same family of perceptible shapes in the same ways. In this case, the family of visual shapes\*

---

<sup>78</sup> One might argue that the shapes\* of tactile sensations could be the same as those of visual sensations. After all, the same perceptible shapes of objects can be determined by the boundaries of both colors and textures. However, if visual shapes\* are dependent on colors\*, and tactile shapes\* are dependent on textures\*, and if we have no reason to conclude that visual shapes\* are dependent on textures\* or that tactile shapes\* are dependent on colors\*, then we have no reason at this point to think that visual and tactile shapes\* are the same.

and the family of tactile shapes\* would be isomorphic to each other. But two families of properties can be isomorphic without being identical.

### 3. Campbell's Argument for Amodality

But John Campbell (1996a, 1996b) argues that the qualitative characters of seeing and feeling the same shapes are the same. If so, visual and tactile sensations of the same shapes have amodal properties in common, in virtue of which both are sensations of those same shapes. If those properties are amodal, the account of sensing feature conjunctions I argued for in the previous chapter is false, since it is a consequence of that view that shapes\*, the mental qualities of those sensations pertaining to shape, are modality specific.

Campbell's argument rests on his so-called Radical Externalist view of qualitative character, according to which the qualitative properties of perceptual states are constituted by the perceptible properties one perceives. Since one sees and feels the same shapes, Campbell argues, one's visual and tactile perceptions of shape are qualitatively identical with respect to shape. Below I will discuss Campbell's argument for Radical Externalism, and how it leads to his conclusion that visual and tactile sensations of shape are the same. I will then argue that Radical Externalism is too strong, and that Campbell's argument for the amodality of the qualitative character of shape perception fails.

Campbell argues that to account for the qualitative character of shape perception we must explain how we perceive the so-called categorical shapes of

objects. Campbell argues, against Sydney Shoemaker's (1984) general theory of properties, that the shapes we sense are not merely conditional, causal properties of objects in virtue of which they behave in certain ways, but the categorical grounds of those conditional powers. According to Campbell, a theory of the qualitative character of sensing shape must account for our perception of such categorical shapes.

According to Shoemaker, properties are conditional, causal powers in virtue of which objects behave in the ways they do. On this view, properties are theoretical posits posited to explain the behavior of objects. If all properties are such conditional powers, the shapes of objects are conditional powers. Accordingly, being spherical, e.g., is a property in virtue of which something will roll down certain inclines and plug holes that have diameters smaller than that of the object, provided that certain other conditions occur. For example, a spherical object will roll down an incline only if the object and the incline are rigid enough, only if there is enough but not too much friction between the object and the incline, and only if no force greater or equal to the force pushing the object down the incline opposes that force. Shapes, on this view, can be exhaustively specified by Ramsey sentences describing such causal roles.<sup>79</sup>

---

<sup>79</sup> The property that realizes the shape specified by such a Ramsey sentence could of course be a categorical property, i.e., a property specifiable independent of that Ramsey sentence. However, according to Shoemaker, what it is for something to have a certain shape is for it to satisfy the Ramsey sentence specifying that shape.

If shapes are such purely causal, conditional powers, then one perceives the shape of an object by perceiving that the object has such a causal power.

But, Campbell argues, in addition to perceiving objects as having such causal powers, we also perceive the categorical grounds of those powers, i.e., the properties in virtue of which an object has the causal power specified by a Ramsey sentence describing a shape.

According to Campbell, our pretheoretic intuitions suggest that we perceive more than just the conditional powers of objects when we perceive the shapes of those objects. Campbell writes, “On the face of it, we do not perceive the shape of a thing as a collection of unsubstantiated threats and promises as to which powers it will take on in various hypothetical circumstances. We perceive the substance behind the threats and promises” (1996a, p. 306).<sup>80</sup> If Campbell is right that we do in fact take ourselves to perceive categorical shapes, not just the conditional powers of shapes, we must explain why we take ourselves to do so. Perhaps the best explanation of this intuition is that we do in fact perceive categorical shapes.

Campbell also argues that the view that we perceive categorical shapes best explains the systematic relations between the appearances of shapes, and it best explains how those systematic relations help account for why shapes have the causal roles they have (1996a, p. 313). One can roll round objects, e.g., but

---

<sup>80</sup> It isn't clear why Campbell claims we do not perceive shapes as causal powers. We could perceive shapes as both causal powers and as categorical grounds of those causal powers.

one cannot roll objects with polygonal cross sections, e.g., cubes and pyramids. And the appearances of round objects, such as spheres and eggs, are more similar to each other than they are to objects with polygonal cross sections. So there is a correlation between the appearances of shapes and the causal roles of shapes. That the appearances of shapes resemble and differ from each other in ways that help account for their causal roles suggests we perceive more than just the causal, conditional powers of shapes.

If Campbell is right that we perceive categorical shapes, we must explain the qualitative character of doing so to fully explain the qualitative character of sensing shapes.

Perhaps we can explain the qualitative character of sensing categorical shape in terms of the view that qualitative character is determined solely by the way we are conscious of our sensations, e.g., introspectively. On this widely held view of qualitative character, the qualitative character of a sensation is determined by what it's like for one to have that sensation. As Campbell characterizes this view, which he calls internalism and attributes to Peter Strawson (1966), "... sensations of shape, and indeed all perceptual experiences, are stratified into similarity classes prior to any environmental circumstances coming into play: they are intrinsically more or less like one another in this or that respect, such as experiential shape or colour" (1996a, p. 302). According to internalism, the sensation in virtue of which one sees a square, e.g., is to be individuated in respect of the similarities and differences

one is conscious of that sensation as bearing to other sensations of shape; i.e., one is conscious of that sensation as resembling the sensation one has when one sees a rectangle more than the sensation one has when one sees a triangle.

Campbell does not explain how he thinks an internalist would attempt to account for our perception of categorical shapes. Perhaps, he might think, internalism could account for categorical shape perception by accounting for the systematic relations among the appearances of shapes. The way a shape appears to one depends in part on the perceptual states one has when perceiving it, i.e., one's sensations of it. So, if the sensations one has when one sees shapes are systematically related to each other as internalism holds, perhaps those relations could account for the systematic relations among the appearances of shapes.

However, if the sensations in virtue of which we sense shapes are individuated internally, only in respect of the similarities and differences revealed in one's awareness of one's own sensations, then those sensations are not individuated in respect of their relations to the physical shapes of objects. If the systematic relations between one's sensations are determined internally, it is not clear how they could make one aware of the systematic relations among physical shapes. If they do not make one aware of such relations, they do not enable one to perceive categorical shapes.

Perhaps internalism could account for one's perceiving shapes as categorical, even if they do not in fact enable one to perceive categorical shapes.

It could be that when one sees a shape, one has a visual sensation of that shape, and one mistakes the properties of that sensation for the shape one sees. Since the sensation is individuated in respect of the ways it resembles and differs from other such sensations, its properties are categorical, not merely conditional, causal properties. So, if one mistook the properties of the sensations one has when one perceives shapes as properties of stimuli, one could perceive those stimuli as having the categorical properties of the sensations. On this view, internalism is a Lockean projectivist, error theory.<sup>81</sup>

If internalism is correct, the qualitative characters of seeing and feeling the same shapes could differ. Visual and tactile sensations of the same shapes seem subjectively different from each other. If such sensations are to be individuated solely by how they seem to one from one's own point of view, i.e., in terms of how we are conscious of them, the mental qualities of sensations pertaining to shape in sight and touch are arguably different.

But perhaps such differences in the way we are conscious of visual and tactile sensations of the same shape result from differences between mental qualities other than those pertaining to shape.<sup>82</sup> Visual sensations of shape have mental qualities pertaining not only to objects' shapes but also to their colors, whereas tactile sensations of shape do not. And tactile sensations of shape

---

<sup>81</sup> The view that we systematically err whenever we perceive shape is problematic in itself. But I won't go into the problems with such theories here (see Meehan, 2003).

<sup>82</sup> See Campbell (1996b, p. 357) and Dretske (1994, p. 95) for similar claims.

have mental qualities pertaining not only to objects' shapes but also to their textures, temperatures, and resistance, whereas visual sensations of shape do not. Perhaps those modality-specific mental qualities determine the introspectible differences between visual and tactile sensations of the same shapes, but the mental qualities of those sensations pertaining to shape are the same.

Internalism is thus compatible with both the view that mental qualities pertaining to shape are modality specific and the view that they are amodal.

One might think that it is a benefit of internalism that it is compatible with both the view that shapes\* are amodal and the view that shapes\* are modality specific. However, this compatibility in fact poses a problem for internalism. According to internalism, the qualitative character of a sensation is determined solely by how one is conscious of that sensation. On this view, one is conscious of all aspects of the qualitative characters of one's sensations. Accordingly, if one is conscious of one's sensation as having a particular mental quality Q, one's sensation has Q, and if one is not conscious of one's sensation as having Q, one's sensation does not have Q. So, if visual and tactile sensations of the same shapes have amodal shapes\* in common, in virtue of which they are sensations of that same shape, one would be conscious that those sensations have the same shapes\*. And, if visual and tactile sensations of the same shapes have different, modality-specific shapes\*, one would be conscious that they do. So, if internalism were correct, it would be obvious to us whether visual and tactile

sensations of shape have amodal shapes\* in common. But it is not at all obvious to us whether the shapes\* of visual and tactile sensations are the same. So internalism is false.

Further, as Campbell argues (1996a, p. 303), internalism runs afoul of Wittgenstein's (1953) private-language argument.<sup>83</sup> If the properties of one's sensations of shape are determined solely by how one is conscious of one's sensations, it is unclear how we would be able to determine whether two people have the same sensations when, e.g., seeing the same shape. If one's sensations are accessible only from one's own point of view, one could not determine whether another perceiver who reports having a visual sensation of a square is reporting the same kind of state one has when seeing a square, since one could not determine whether the other perceiver uses the expression 'sensation of a square' to refer to the same state as one does when one utters that expression.

Further, if one can access sensations only from one's own point of view, it is unclear how one would be able to determine whether one has the same sensation when seeing the same shape on different occasions. One might argue that one could simply remember the sensation one had previously and compare it with one's current sensation. But, if first-person access to one's own sensations provides the only access to one's sensations, one could not

---

<sup>83</sup> Though an exhaustive discussion of the private-language argument and the controversy surrounding it is well beyond the scope of this dissertation, I will briefly defend what I take to be Campbell's interpretation.

determine whether one is correctly remembering the sensation one had when one saw the shape on another occasion. Memory is fallible, and we often consult external sources to determine whether we are remembering something correctly, e.g., when asking someone else to describe an event we are trying to remember. But, if one is one's only source of information about one's own mental states, there is no source other than one's own memory one could consult to determine whether one is correctly remembering the sensation one had when seeing a shape on some other occasion. So the view that qualitative character is determined solely by how one is conscious of one's sensations provides no way to determine whether one is remembering one's sensation correctly. So it provides no way to determine whether one has the same kind of sensation when seeing the same shape on different occasions.

But presumably we can determine these things. As Dan Dennett (2005, pp. 30-31) claims, we presuppose that we can do so whenever communicating with each other about sensations. If we can in fact determine these things, the mental qualities of sensations pertaining to shape, and to all other perceptible properties, are not determined solely by how one's sensations seem to one; rather, we have intersubjective access to sensations. So we must explain sensations and their properties in a way that accounts for our intersubjective, third-person access to them, as well as our diachronic first-person access to them.

However, following Saul Kripke (1982), one might argue that Wittgenstein's private-language argument poses a skeptical puzzle along the lines of Goodman's (1953) new riddle of induction. If so, perhaps the private-language argument poses no more threat to internalism about qualitative character than Goodman's riddle poses to our application of color predicates or Hume's riddle of induction poses to causal explanations.

Kripke claims that the private-language argument is an instance of the more general problem of rule following. Rules apply to an indefinite number of cases. However, no rule has been applied more than a finite number of times. This poses a problem for one's justification in applying a particular rule. When one applies a rule, one does so on the basis of its past applications. But there is no way for one to know what rule to apply in a new case, since there is no way for one to know what rule one applied in those earlier cases.

Kripke uses an example from arithmetic to illustrate the paradox. Suppose one has never added 68 and 57 and one is given the task of doing so. One easily calculates the sum 125. To perform this calculation, one arguably applies rules of arithmetic one has applied on many other occasions, e.g., when adding 5 and 6, 25 and 32, and 3 and 4. But, a skeptic might argue, one is not justified in calculating 125 rather than some other answer, e.g., 5. According to the skeptic, it is logically possible that the correct rule to follow in this case is not one that generates the answer of 125.

One might object that one is simply applying the same rules of arithmetic one has applied on all other occasions, and that those prior applications justify one's answer to the new arithmetical problem. But, the skeptic could argue, one does not know what rule one applied on those prior occasions; again, rules apply to an indefinite number of cases, but one has performed a finite number of arithmetical problems. So, perhaps, the rule one followed in the past gives us 5, not 125, as the sum of 68 and 57.

Kripke claims that Wittgenstein's private-language argument provides another example of this same skeptical puzzle. One has made only a finite number of first-person ascriptions of sensations. And to ascribe a sensation to oneself now, one must follow some rule governing the application of sensation predicates, e.g., 'has a sensation of a red square'. But, the skeptic might argue, one has no way to know what rule to follow in this case. One cannot simply appeal to the rules one followed in the past, since one cannot know what rules one applied in the past. So when one says, "I'm having a sensation of a red square," one cannot know for sure whether one is referring to the same kind of sensation now that one referred to in the past when one also said, "I'm having a sensation of a red square."

If one interprets the private-language argument in this way, one might claim that it fails to undermine internalism about qualitative character. Even if we cannot know for sure what rules we follow when ascribing sensations to ourselves on different occasions, that does not show that we fail to do so

accurately, it just shows that we do not know for sure if we are doing so accurately. Nevertheless, there are ways in which we can proceed in making accurate ascriptions of sensations.

But Kripke ignores an important difference between one's referring to sensations and one's doing arithmetic. There are a number of ways to check one's arithmetic; one can consult other people, textbooks, a calculator, or an abacus. However, if we have only first-person access to our sensations, there are no ways to check whether we are accurately referring to them, nor whether we are referring to them in the same ways on different occasions. Wittgenstein is concerned with internalism's failure to allow for such independent checks on our ascriptions of sensations. This problem does not arise for arithmetic, since one's calculations are publicly observable. So the problem Wittgenstein raises for internalism about qualitative character is not the same as the problem he and Kripke raise for rule following; the problem Wittgenstein raises against internalism poses a more serious challenge.

So we must account for the qualitative character of sensations in a way that allows for such checks on our ascriptions of sensations. We must account for intersubjective and diachronic access to sensations.

Campbell argues that we can account for such access to sensations with the view that sensations are to be individuated in respect of the physical properties one perceives. According to Campbell, "... the sorting of sensations into similarity classes constitutively demands an appeal to the environment..."

(1996a, p. 303), and "[t]he geometrical aspects of one's experience of objects will then be constituted by the geometry of the objects in one's surroundings" (1996a, p. 302). On this view, which Campbell calls radical externalism, the qualitative character of seeing a shape is determined by the shape of the object one sees,<sup>84</sup> and the difference between the qualitative characters of seeing, e.g., a square and seeing a triangle is determined by the differences between the shapes of the objects one sees.

If qualitative character is determined by the properties of the objects one senses, it is intersubjectively accessible, since those properties are perceptible by multiple perceivers. So Campbell's radical externalism accounts for how one determines whether one has the same sensation on different occasions and whether two people sensing the same stimuli have the same sensations.

Campbell claims it is a consequence of radical externalism that the qualitative characters of seeing and feeling the same shapes are themselves the same. If the qualitative character of seeing a square is constituted by the shape one sees, and the qualitative character of feeling a square is constituted by the

---

<sup>84</sup> Campbell's view is similar in this respect to Aristotle's view that one's perceptions have the same properties as the stimuli one senses, e.g., sensations of red squares are themselves red and square (*de Anima* II, 5, 418a4, II, 11, 423b31, II, 12, 424a18, III, 2, 425b23). Also, Campbell (2002) argues for the Russellian view that the objects we perceive are themselves constituents of one's perceptions, and that it is because those objects are constituents of one's perceptions that one knows what one is referring to or thinking about. But Campbell does not discuss the qualitative character of seeing and feeling the same shape there. Campbell (2005) further develops his view of perceptual experience, but he remains agnostic about whether the qualitative characters of seeing and feeling the same shapes are themselves the same.

shape one feels, then the qualitative characters of seeing and feeling a square are constituted by the same physical shape. So, Campbell argues, seeing and feeling the same shape have the same qualitative character (1996a, p. 303). If so, the properties of visual and tactile sensations pertaining to shape are the same, they are amodal.

#### 4. An Objection to Amodality

But the qualitative properties of visual and tactile shapes are doubtless not constituted entirely by the shapes one perceives, even if we must appeal to those shapes to account for qualitative character. The properties of visual and tactile sensations pertaining to shape are presumably determined in part by properties of the visual and tactile perceptual systems.<sup>85</sup> And differences between those perceptual systems likely contribute to differences between the properties of visual and tactile sensations of shape. So, even if sensations of shape are to be individuated partly in respect of the shapes we sense, the properties of those sensations arguably differ in sight and touch.

#### 5. Homomorphism Theory and Modality Specificity

So a theory of the qualitative character of sensing shape must allow for differences between the mental qualities of visual and tactile sensations

---

<sup>85</sup> Brian Loar briefly raises this point (1996, p. 322).

pertaining to shape, while also accounting for third-person access to sensations. Both internalism and Campbell's radical externalism fail to do so.<sup>86</sup>

We can explain how we sense shape while meeting both requirements in terms of homomorphism theory, the view that mental qualities represent perceptible properties in virtue of resembling and differing from each other in ways parallel to the ways perceptible properties resemble and differ from each other (Rosenthal, 1991, 2005; Meehan, 2002, 2003). According to homomorphism theory, the visual mental quality pertaining to physical, perceptible squares, e.g., resembles and differs from other mental qualities pertaining to shape in ways parallel to the ways physical, perceptible squares resemble and differ from other perceptible shapes. Just as squares are more similar to rectangles than triangles, the mental quality pertaining to squares is more similar to that pertaining to rectangles than that pertaining to triangles.

Homomorphism theory explains more than just the introspectible similarities and differences between sensations. According to homomorphism theory, mental qualities are theoretical posits, posited to explain how we discriminate among perceptible properties, e.g., shapes. It is because sensations have mental qualities that resemble and differ from each other in ways parallel to the ways shapes resemble and differ from each other that we can discriminate shapes on the basis of their similarities and differences. The

---

<sup>86</sup> Of course, a theory of qualitative character need not affirm the view that the mental qualities of visual and tactile sensations of shape are different. Rather, such a theory must not rule out the modality specificity of such mental qualities.

similarities and differences between mental qualities pertaining to shape enable us to see squares, e.g., as more similar to rectangles than triangles, and they enable us to feel squares as more similar to rectangles than triangles.

Since the sensory discriminations one makes are observable, homomorphism theory accounts for third-person access to sensations. Two people who visually discriminate among shapes in the same ways have the same visual sensations of shape.<sup>87</sup> And if one makes the same visual shape discriminations on different occasions, one has the same visual sensations of shape on those occasions. So homomorphism theory, unlike internalism, does not run afoul of Wittgenstein's private-language argument.

But homomorphism theory, unlike Campbell's radical externalism, allows for differences between the mental qualities of visual and tactile sensations pertaining to shape. According to homomorphism theory, mental qualities of both visual and tactile sensations represent shape in virtue of resembling and differing in ways parallel to the ways shapes resemble and differ from each other. But it could be that the visual mental qualities pertaining to shape resemble and differ

---

<sup>87</sup> If sensations had intrinsic, nonrepresentational properties, as the internalist argues, two people could perhaps have intrinsically different sensations while making the same shape discriminations. Those same shape discriminations, the internalist could argue, would be underwritten by the same representational properties of those intrinsically distinct sensations. But if we can account for the qualitative character of sensing shapes without adverting to such intrinsic properties of sensations, it is unclear why one would think that sensations have those intrinsic properties. I am arguing here, as I have argued in chapter 1, that appeal to such intrinsic properties leads to problematic commitments, and that we need not appeal to intrinsic properties of sensations in order to explain qualitative character.

from each other in ways distinct from the ways tactile mental qualities pertaining to shape do; so they could represent physical shapes in respect of different relations of similarity and difference. So homomorphism theory allows for the modality specificity of mental qualities pertaining to shape.

## 6. Crossmodal Transfer of Shape Information

Experiments on our ability to recognize shape across sensory modalities could also provide insight into the relations between the mental qualities of sensations pertaining to shape in different modalities. Such experiments examine one's ability to recognize in one sensory modality a stimulus or stimulus property one previously sensed only in another modality, e.g., when one visually recognizes a shape one felt but did not see earlier.

Such crossmodal recognition is seemingly automatic and effortless; one rarely has trouble, e.g., visually recognizing a shape one has previously felt. Perhaps, one might argue, crossmodal shape recognition is so effortless because the sensations in virtue of which one sees and feels shape have some amodal property in common, in virtue of which they are sensations of the same shape. Accordingly, one visually recognizes the triangle one previously felt because one's visual sensation of the triangle and one's prior tactile sensation of the triangle share some amodal mental quality, i.e., an amodal triangular\* shape\*.

On the other hand, one could arguably perform crossmodal transfer of shape information even if visual and tactile sensations of shape do not share a common, amodal property in virtue of which both sensations are sensations of that same shape. As the empiricists Berkeley (1732/1975), Locke, and Molyneux (see Locke, 1690/1975, p. 146) argued, correlations between sensations of the same shapes in sight and touch could be learned. If correlations between visual and tactile sensations of shape are in fact learned, crossmodal transfer rests on correlations between distinct mental qualities of visual and tactile sensations pertaining to the same shapes. And we can explain the ease with which one performs crossmodal transfer of shape information without committing to the view that visual and tactile sensations of the same shape have some amodal property in common. A great deal of effortless, seemingly automatic behavior is underwritten by processes that coordinate distinct mental states. For example, adults read text written in their native language with great ease. But reading involves a process whereby visual information about the shapes and spatial relations of letters is correlated with semantic information; the visual sensations of words and letters in one's native language are distinct from the intentional states involved in one's comprehension of the text. Likewise, understanding another speaker's utterances involves a process by which auditory information is correlated with semantic information. And reading Braille involves a process by which tactile information is correlated with semantic information; so tactile sensations must be correlated with the intentional states in virtue of which one

understands the Braille text. Nevertheless, these cases of language comprehension are seemingly automatic to proficient readers and speakers of a language and to those who proficiently read Braille. So the ease with which one performs crossmodal shape integration does not by itself show that the properties of visual and tactile sensations pertaining to the same shapes are themselves the same. Crossmodal transfer of shape information could rest on a process that correlates distinct, modality-specific shapes\*, even if it is seemingly automatic.

And, if the mental qualities pertaining to shape are in fact determined by the boundaries of other, modality-specific mental qualities, as I've argued, then the mental qualities of visual and tactile sensations pertaining to the same shapes are themselves different. If so, visual and tactile sensations of the same shapes must be correlated to enable crossmodal transfer of shape information.

Data showing that the ability to transfer shape information across modalities is learned would further support the claim that the mental qualities of visual and tactile sensations pertaining to shape are distinct. If the mental qualities of visual and tactile sensations pertaining to shape were the same, crossmodal transfer would occur automatically, since there would be no relevant difference between visual and tactile sensations of shape one would need to correlate. On the other hand, if we do not learn to transfer shape information across modalities, i.e., if that ability is innate, it could be that visual and tactile

mental qualities pertaining to shape are the same, or that they are different but innately correlated.

In fact, a great deal of research suggests that infants as young as 16-hours old can perform crossmodal transfer tasks, both between vision and touch and between vision and proprioception. If newborn infants can perform crossmodal transfer of shape information, one might argue, crossmodal transfer does not rest on learned correlations between sensations with only modality-specific mental qualities.<sup>88</sup> Rather, one might argue, the ability of infants to transfer spatial information across different sensory modalities reflects an innate ability to transfer such information across sensory modalities.

One might further argue that if crossmodal transfer abilities are innate, they are best explained in terms of the view that sensations of the same stimulus properties in different sensory modalities have amodal mental qualities representing those stimulus properties in common.<sup>89</sup>

---

<sup>88</sup> I am not claiming that sensations in different modalities have no amodal properties in common, just that visual and tactile sensations of the same perceptible spatial properties do not have mental qualities in common, in virtue of which they are sensations of the same perceptible spatial properties. Visual and tactile sensations of a square do not have the same mental shape\* qualities, but they could have the same mental qualities pertaining to temporal properties, such as those pertaining to stimulus onset or order.

<sup>89</sup> Meltzoff (1993) claims experiments showing that infants transfer shape information across modalities show that shape perception involves a supramodal perceptual system, and Bermúdez (1998) claims these experiments show that shape perception is not modality specific. However, it's not entirely clear how these claims bear on the question of the relations between mental qualities in different modalities. But they could reasonably be taken as the claim that the mental qualities pertaining to shape are the same in sight and touch.

I'll discuss three forms of crossmodal-transfer experiments. These include experiments on the abilities of infants to imitate facial expressions, to visually recognize shapes they have felt orally, and to visually recognize shapes they have felt manually. I'll argue that none of these forms of crossmodal transfer shows that sensations of spatial properties in different sensory modalities have amodal mental qualities in common, nor do they show that crossmodal transfer is innate.

Finally, I'll argue that asymmetries in crossmodal shape transfer during different stages of development suggest that the properties of visual and tactile sensations pertaining to the same shapes are different.

### 7. Crossmodal Transfer in Infants: Facial Imitation

Andrew Meltzoff and Keith Moore (1977, 1983) tested newborn infants, 2-3 weeks old and 3 days old and younger, respectively, to determine whether they could imitate the facial expressions of others. If infants can in fact imitate facial expressions, their ability to do so must be innate, not learned, since newborn infants have no visual perception of their own facial expressions in virtue of which they could determine whether they are moving their facial muscles in the right way to produce the facial expressions they see. So, perhaps, to imitate facial expressions, infants must match their proprioceptive and kinesthetic sensations of the form and movements of their own facial features with the form and

---

movements of the facial features of the model they see. If they can do this, then either the properties of visual, proprioceptive, and kinesthetic sensations pertaining to the form and movements of facial features are the same, or those properties are distinct but innately coordinated.

Meltzoff and Moore (1983) tested whether newborn infants could imitate simple facial expressions, such as protruding one's tongue or opening one's mouth. The infants watched an adult model produce one of these expressions for 20 seconds. While watching the adults, the infants sucked on a pacifier. The adult then stopped producing the facial expression, and the pacifier was removed from the infant's mouth for 20 seconds. They ran six such trials on each infant.

Meltzoff and Moore found that infants produced significantly more tongue protrusions following the model's tongue protrusions than following the model's mouth openings. And they found that infants produced significantly more mouth openings following the model's mouth openings than following the model's tongue protrusions. Meltzoff and Moore claim this shows a strong correlation between infants' visually observing facial expressions and infants' producing matching facial expressions. So, they conclude, newborn infants successfully imitate adults' simple facial expressions.

They further conclude that, since the infants could not have learned to correlate their visual perceptions of the adults' facial expressions and the feelings of their own facial expressions, such correlations are innate and involve a supramodal perceptual system.

However, Moshe Anisfeld (1991, 1996) argues that Meltzoff and Moore's experiments show a correlation only between infants' seeing tongue protrusions and their producing tongue protrusions, they do not show a correlation between infants' seeing mouth openings and producing mouth openings. If infants match only a single facial expression, Anisfeld argues, that matching behavior is best explained in terms of an innate releasing mechanism, not a supramodal perceptual system in virtue of which one can imitate facial expressions. If infants could imitate facial expressions, she claims, such imitation would not be limited to a single expression.

According to Anisfeld, Meltzoff and Moore's conclusion that infants imitate mouth openings rests on a statistical confound. Meltzoff and Moore concluded that infants imitate mouth openings because infants produced more mouth openings after seeing mouth openings than they did after seeing tongue protrusions. However, if there is a strong correlation between infants' seeing tongue protrusions and their producing tongue protrusions, they would produce relatively few mouth openings after seeing tongue protrusions; since the infants are producing tongue protrusions, they cannot produce mouth openings. So it could be that infants produce more mouth openings after seeing mouth openings than after seeing tongue protrusions simply because their producing a significant number of tongue protrusions after seeing tongue protrusions prohibits them from producing mouth openings after seeing tongue protrusion, not because they produce more mouth openings after seeing mouth openings. So, Anisfeld

claims, the appearance of infants' ability to imitate mouth openings could be due to the significant correlation between their seeing and producing tongue protrusions. So Meltzoff and Moore fail to show that infants imitate both tongue protrusions and mouth openings.

In fact, though the infants in Meltzoff and Moore's experiment produced significantly more tongue protrusions than mouth openings after seeing tongue protrusions, they did not produce significantly more mouth openings than tongue protrusions after seeing mouth openings. This further supports Anisfeld's interpretation of their data. If infants imitate mouth openings after seeing them, they would produce more mouth openings than any other expression after seeing the adult models produce mouth openings.

Anisfeld, Gerald Turkewitz, and Susan Rose (2001) replicated Meltzoff and Moore's experiment to test Anisfeld's hypothesis. However, in addition to monitoring the infants' behavior during and after they watched adult models produce tongue protrusions and mouth openings, Anisfeld et al. also included control trials in which the adult models produced no facial expressions at all.

Like Meltzoff and Moore, Anisfeld et al. found that the infants produced significantly more tongue protrusions than mouth openings after seeing the model produce a tongue protrusion. But they found that infants produced the same number of mouth openings after seeing mouth openings as they did after seeing tongue protrusions. This data supports Anisfeld's claim that Meltzoff and Moore failed to show that infants imitate mouth openings.

Further, Anisfeld et al. found that infants in the control condition, the condition in which the adult model produced no facial expression, produced more tongue protrusions than mouth openings. And infants produced no more tongue protrusions after seeing tongue protrusions and no more mouth openings after seeing mouth openings than they did in the control condition. This suggests that the infants' tongue protrusions reflect something other than an ability to imitate such expressions.

Susan Jones (1996) argues that infants stick out their tongues not only in response to adults' doing so, but also in response to interesting, novel stimuli. Jones showed that infants who show interest in certain light displays by staring longer at them than they stare at other stimuli also produce more tongue protrusions when seeing those light displays than infants who do not stare longer at those light displays than at other stimuli. She concludes that tongue protrusion could be a reaction to interesting stimuli; perhaps infants stick out their tongues to tactually explore interesting visual stimuli. It could be that the infants in Meltzoff and Moore's experiments stick out their tongues because they are interested in the stimuli they see, not because they are imitating the expressions of the adult models. If so, Meltzoff and Moore fail to show that infants can transfer spatial information about facial expressions from vision to proprioception and kinesthesia.

But, even if the infants in Meltzoff and Moore's experiments did in fact imitate the expressions of the adult models, this would not show that

proprioceptive and kinesthetic sensations of the spatial configurations of facial features have amodal mental qualities in common with visual sensations of those expressions. It could be that the mental qualities of a visual sensation pertaining to the position, shape, and movement of a tongue protruding from someone's mouth are different from the mental qualities of one's proprioceptive and kinesthetic sensations pertaining to the position, movement, and shape of one's own tongue, but the correlations between those modality-specific mental qualities that are needed to perform visual-to-proprioceptive/kinesthetic transfer of spatial information are innate. Since the view that the mental qualities of visual and proprioceptive and kinesthetic sensations pertaining to the same spatial properties of facial expressions are themselves different is compatible with the view that visual-to-proprioceptive/kinesthetic transfer of such spatial information is innate, Meltzoff and Moore's experiments do not undermine modality specificity, even if they succeed in showing that the ability to imitate facial expressions is innate.

Further, it could be that such imitative abilities do not involve sensations. As I discussed in chapter 2, there is a good deal of evidence for the existence of two separate visual processing streams, one underlying visual perception and the other underlying visuomotor action (Goodale and Murphy, 1997; Haffenden, Schiff, and Goodale, 2001; Milner and Goodale, 1995; Perenin and Vighetto, 1988). If the states involved in visual perception and those involved in visuomotor action are distinct, it could be that the former but not the latter are

sensations. Visuomotor action could involve the correlation of subpersonal motor codes that direct bodily movements and subpersonal visual states that encode information about visual stimuli. If so, then even if those visual and motor states have properties in common, that does not show that visual sensations and proprioceptive and kinesthetic sensations have mental qualities in common. It could be that the states underlying infants' imitative abilities, if they do in fact have those abilities, are such subpersonal states, not sensations.

#### 8. Tactile-to-Visual Shape Transfer

Meltzoff and Borton (1979) showed that one-month-old infants look longer at shapes they have previously felt but not seen than they look at novel shapes, i.e., shapes they have not previously felt or seen. This differential treatment of novel and previously felt shapes suggests that infants prefer to look at familiar shapes, suggesting in turn that they visually recognize shapes they have felt before. If so, these experiments show one-month-old infants perform crossmodal transfer of shape information from touch to vision.

The infants in Meltzoff and Borton's experiment first orally explored one of two pacifiers, either a smooth pacifier or a pacifier with a number of small protuberances, or nubs, on it. They explored the pacifier for 90 seconds without seeing it. Then, once the pacifier was removed, the infants were shown two objects that they were prevented from feeling. One of these objects was the shape of the smooth pacifier and the other was the shape of the nubbed pacifier.

Meltzoff and Borton found that infants who had orally explored the smooth pacifier subsequently looked longer at the smooth visual stimulus, whereas infants who had orally explored the nubbed pacifier subsequently looked longer at the nubbed visual stimulus, suggesting that they are able to transfer information about the shapes of stimuli from touch to sight. Meltzoff and Borton further claim the infants have not had sufficient oral-tactile and visual experience of these shapes to learn to transfer shape information from touch to shape. So, they conclude, the infants' ability to transfer shape information from touch to sight is innate.

But it could be that the infants have in fact had enough experience with these shapes to coordinate tactile and visual sensations of them. The smooth pacifier is similar in shape to the infants' own fingertips, their mother's nipples, and the nipples on feeding bottles (Rosenthal, 2005). And infants have considerable visual and tactile experience of their own fingers, their mothers' nipples, and the nipples on bottles during their first month of life.

The nubs on the nubbed pacifier are also similar in shape to that of the infants' fingertips, their mother's nipples, and the nipples on bottles. So the infants could exploit the same correlations between visual and tactile sensations of shape when they visually recognize a nubbed visual stimulus after orally exploring the nubbed pacifier as they do when they visually recognize the smooth visual stimulus after orally exploring the smooth pacifier. The smooth tactile and visual stimuli feel and look different from the nubbed stimuli because the nubbed

stimuli have several small protrusions on them. But those protrusions are themselves similar in shape to the shapes of the infants' fingertips, their mothers' nipples, and the nipples on a bottle, just as the smooth pacifier and smooth visual stimulus are. So, once an infant has learned to correlate tactile and visual sensations of nipple-shaped objects, the infant can perform Meltzoff and Borton's crossmodal-transfer task.

According to Meltzoff, the infants could not have learned to correlate visual and tactile sensations of the shape of a nipple, since they do not, e.g., look at their mother's nipples while feeding (1993, pp. 223-224). But the infants could draw correlations between their tactile and visual sensations of this shape without ever simultaneously seeing and feeling it. Repeated visual and tactile exposure to that shape within short time frames could be sufficient for reinforcing correlations between visual and tactile sensations of the shape. And infants do have such exposure to the shape of a nipple whenever they feed.

Further, other developmental psychologists (Maurer, Stager, and Mondloch, 1996; Brown and Gottfried, 1986; Pêcheux, Lepecq, and Salzarulo, 1988) attempted to replicate Meltzoff and Borton's results and failed. These failures to replicate Meltzoff and Borton's results call into question Meltzoff and Borton's conclusion that one-month-old infants transfer shape information from touch to sight. If one-month-old infants do not perform tactile-to-visual transfer of shape information, then the ability to do so is arguably learned, presumably at a later stage of development.

However, it could be that these failures to replicate Meltzoff and Borton's results were due to differences between those experiments and Meltzoff and Borton's experiment. For example, as Maurer et al. (1999, p. 1048) note, Borton and Gottfried presented each infant with four pairs of stimuli, not just the smooth and nubby pair. So their task was presumably more difficult than Meltzoff and Borton's task. Also, it could be that Borton and Gottfried's subjects did in fact perform crossmodal transfer with the smooth and nubby stimuli, but the evidence of that crossmodal transfer was obscured by Borton and Gottfried's collapsing the data across all four pairs of stimuli before analyzing it. Finally, the nubs on the nubbed stimuli Pêcheux et al. presented to their subjects were significantly smaller than those in Meltzoff and Borton's experiment. So it could be that the infants in the Pêcheux et al. experiment failed to feel or see the difference between the smooth and nubby stimuli.

On the other hand, perhaps the infants in Meltzoff and Borton's experiment looked longer at the shapes they had orally explored, not because they recognized those shapes, but for some other reason. Daphne Maurer, Christine Stager, and Catherine Mondloch (1996) argue that we can best explain Meltzoff and Borton's results in terms of infants' preferences to look at objects that are located off to one side in their visual fields. If an infant is visually biased towards looking to the left, e.g., and the matching visual stimulus is presented on the infants' left side, the infant will look longer at that stimulus than at the nonmatching stimulus presented on the right side. In such a case, it could be

that the infant looks longer at the matching stimulus, not because it matches the tactile stimulus the infant previously explored, but because it is presented on the side the infant prefers to look at.

Meltzoff and Borton did attempt to control for such side biases by presenting the matching visual stimuli on the left to half of the infants and on the right to the other half. But, according to Maurer et al., this method is inadequate for a sample size as small as Meltzoff and Borton's. Since Meltzoff and Borton tested only 32 infants, Maurer et al. argue, it could be that the matching shape was inadvertently presented to more than half of the subjects on their biased side. If they presented the matching shape on the biased side for a significant number of the infants, then those infants would look longer at the matching shape than at the novel shape. But, in this case, it would be unclear whether the infants looked at the matching shape because it matched the shape they were habituated to, rather than because it happened to appear on their biased side. If so, it could be that Meltzoff and Borton's results are due to infants' visual side biases, not their ability to transfer shape information from touch to sight.

To test whether Meltzoff and Borton's results were in fact due to such visual side biases, as opposed to crossmodal transfer of shape information, Maurer et al. replicated their original experiment with new controls for side bias. Instead of presenting the matching stimulus on the left to half the subjects and on the right to the other half, they ran two trials for each infant. In the first trial, they presented the matching stimulus on one side, and on the second trial, they

presented the matching stimulus on the other side. If Meltzoff and Borton's results were due to infants' crossmodal transfer of shape information, Maurer et al. reasoned, switching the sides on which the matching and nonmatching stimuli are presented will not affect infants' looking times.

Maurer et al. found that about half of the infants showed a strong side bias; they looked to one side for a mean of 80% of the time during both trials. Of these infants with side biases, half preferred looking to the left, and half to the right.

When analyzing the data across all of the infants' first trials, they found no evidence of crossmodal transfer of shape information from touch to sight. The subjects orally habituated to a pacifier looked longer at the matching visual stimulus only 43.5% of the time. Further, only 12 of the 32 subjects looked longer at the stimulus that matched the shape of the pacifier they had orally explored.

But, though there was no evidence of crossmodal transfer of shape information when all of the data were analyzed together, the data show that infants habituated to the different pacifiers behaved differently when presented with the visual stimuli. Infants orally habituated to the nubby pacifier did not show a preference for one shape over the other; 10 subjects looked longer at the nubby pacifier, and 6 looked longer at the smooth pacifier. According to Maurer et al., this distribution is not significantly different from chance. However, infants orally habituated to the smooth pacifier looked at the smooth visual stimulus only

22.6% of the time, which is significantly below chance levels. And only 2 of those 16 subjects looked longer at the smooth visual stimulus in the first trial.

Perhaps, one might argue, the infants habituated to the smooth pacifier performed crossmodal transfer of shape information, but the infants habituated to the nubby pacifier did not.

But there are two problems with this interpretation of the data. The infants in the smooth group looked longer at the nubby visual stimuli, not the smooth visual stimuli. But in Meltzoff and Borton's experiment, infants looked longer at shapes that matched those they had been habituated to; i.e., those habituated to a smooth pacifier looked longer at the smooth visual stimulus, and those habituated to a nubby pacifier looked longer at the nubby visual stimulus. It is unclear how we can account for these opposing results on the view that the infants' looking times in both experiments are due to their transferring shape information from touch to sight.

Also, only the infants habituated to smooth pacifiers showed a looking preference in the Maurer et al. experiment. But both infants habituated to smooth pacifiers and infants habituated to nubby pacifiers showed looking preferences in Meltzoff and Borton's experiment. Again, it is unclear how we can explain this difference on the view that the infants are transferring shape information across modalities.

But we can in fact explain these differences between the Maurer et al. results and Meltzoff and Borton's results in terms of the view that infants' looking

times in these experiments are due to their visual side biases, not to their ability to transfer shape information from touch to sight. Maurer et al. found that "... of the infants who looked longer to the same side across both test trials, the nubby stimulus was placed on the preferred side during trial 1 for 80% of the 10 such infants in the smooth group, but only at a percentage near chance ... for the 11 such infants in the nubby group and 12 such infants in the baseline [i.e., control] group" (1996, p. 1052). So it could be that the infants who had been habituated to the smooth pacifier looked longer at the nubby visual stimulus simply because it was inadvertently presented more often on their preferred sides, whereas no significant coincidence of nonmatching stimulus and preferred side occurred with the visually side-biased infants who had been tactually habituated to the nubby pacifier.

On this view, the infants in the Maurer et al. experiment who were habituated to the smooth pacifier looked longer at the nubby visual stimulus because it happened by chance to appear more often on their biased side, not because the infants visually recognized the smooth visual stimulus and preferred to look at the novel, nubby stimulus. And it is because the nubby pacifier did not appear more often on the infants' biased sides that the infants habituated to the nubby pacifier did not look longer at the smooth pacifier.

This explanation also accounts for the difference in which visual stimulus, i.e., matching or nonmatching, the infants looked at longer in these two experiments. If the infants' looking times are due to their side biases, then

infants presented with the matching visual stimulus on their preferred side will look longer at the matching visual stimulus, and those presented with the nonmatching visual stimulus on their preferred side will look longer at the nonmatching stimulus. It could be that in Meltzoff and Borton's experiment more infants with visual side biases were presented with matching visual stimuli on the side they prefer to look at, while those infants with visual side biases in the Maurer et al. experiment were presented with nonmatching stimuli on the side that they prefer to look at.

Two further experiments support this explanation. In both experiments, the infants Maurer et al. tested showed significantly less visual side bias than the infants in the first experiment. And these infants did not look longer at visual stimuli that matched the shape they had been orally habituated to, nor did they look longer at visual stimuli that did not match the shape they had been orally habituated to; their looking times did not significantly differ when the infants were presented with familiar and novel shapes. This result held equally for infants tactually habituated to the smooth stimulus and those tactually habituated to the nubby stimulus. This provides further evidence of a correlation between infants' looking times and their visual side biases. And it reveals no such correlation between how long infants look at stimuli and the shapes of those stimuli. So these results confirm the view that infants' longer looking times in both Maurer's et al. first experiment and those in Meltzoff and Borton's experiment resulted from infants' visual side biases, not from their recognizing shape across

modalities. Meltzoff and Borton thus fail to show that infants transfer shape information from touch to sight.

So it could be that infants must learn to transfer shape information from touch to sight. And if crossmodal transfer of shape information is learned, then the mental qualities of tactile and visual sensations pertaining to shape are distinct and modality specific.

However, other experiments suggest that newborn infants, as young as 16 hours, can visually recognize shapes that they have only manually explored. Perhaps such experiments show that crossmodal transfer of shape information is in fact innate.<sup>90</sup> Though such a finding would not by itself show that visual and tactile mental qualities representing shapes are the same, it would show that they could be the same, since it would show that crossmodal transfer of shape information is not learned.

#### 9. Crossmodal Transfer in Infants: Visuo-Tactile Shape Transfer

Arlette Streri and Edouard Gentaz (2003) tested newborn infants, with a mean age of 62 hours, to determine whether they could visually recognize shapes they had manually felt but not seen before. Streri and Gentaz assumed that if newborns visually recognize shapes they felt before, this would show that

---

<sup>90</sup> If the experiments I discuss below support the view that crossmodal shape transfer is innate, we must explain the discrepancy between them and the Maurer et al. results I just discussed. However, I'll argue that the experiments on visual-to-tactile shape transfer do not support the view that crossmodal shape transfer is innate.

the ability to transfer shape information across modalities is innate, not learned. Since these infants have not had enough time to learn to correlate felt shapes and seen shapes, or to learn to correlate tactile and visual sensations of the same shapes, Streri and Gentaz argue, any such correlation between vision and touch the infants exhibit is innate.

Streri and Gentaz, like Meltzoff and Borton, used an intersensory paired-preference procedure to determine whether newborns do in fact perform tactile-to-visual shape transfer. Infants were tactually habituated to one of two shapes, either a cylinder or a pyramid, both of which were small enough that the infants could grasp them in one hand.<sup>91</sup> During this tactile habituation stage, the infants were prevented from seeing the objects they were feeling. Once a subject was tactually habituated to the object, the object was taken away, and the subject was shown both a cylinder and a prism, hanging side by side from fishing line.

Streri and Gentaz monitored where subjects looked and recorded how long and how frequently they looked at each object. They found that subjects who had been tactually habituated to a prism looked longer and more frequently at the cylinder, and subjects who had been tactually habituated to a cylinder looked longer and more frequently at the prism. These results reveal that the infants preferred to look at the shape to which they had not been tactually

---

<sup>91</sup> In this experiment, all infants were habituated to the shapes using their right hands. In a subsequent experiment, Streri and Gentaz (2004) found that tactile-to-visual shape transfer occurs for infants' right hands, but not their left hands.

habituated, i.e., the novel object. So, Streri and Gentaz conclude, infants visually recognize the shape they previously felt during the habituation phase of the experiment.<sup>92</sup>

Streri and Gentaz claim that “[t]hese results reveal the ability of newborns to transfer shape information from right hand to eyes before they had the opportunity to learn from the pairing of visual and tactual experience” (2003, p. 17). If so, this suggests that the ability to transfer shape information across modalities is innate, i.e., not learned.

But these experiments do not show that the mental qualities of visual and tactile sensations pertaining to the same shapes are themselves the same. Again, even if the ability to transfer shape information across modalities is in fact innate, it could be that this ability reflects an innate coordination of distinct visual and tactile mental qualities pertaining to the same shapes. On this view, a visual sensation of a cylinder, e.g., has a particular mental quality, visual cylindrical\*, in virtue of which that sensation is a sensation of a cylinder, and a tactile sensation of a cylinder has a different mental quality, tactile cylindrical\*, in virtue of which that sensation is a sensation of a cylinder, but the mechanisms in virtue of which these distinct sensations and mental qualities are correlated are innate, not

---

<sup>92</sup> Streri and Gentaz ran control experiments in which subjects were not tactually habituated to any shape, and found that subjects did not exhibit a visual preference for either shape. They also found that subjects showed no preference for the side on which a visual stimulus was presented; i.e., being habituated with the right hand produces no preference to look at visual stimuli presented on the right side of the display. Finally, Streri and Gentaz (2004) found subjects performed such crossmodal transfer even when the visual stimuli were presented in succession, as opposed to being presented simultaneously.

learned. Again, the view that crossmodal transfer rests on an innate ability is compatible with the view that the mental qualities pertaining to spatial properties, e.g., shapes, are themselves modality specific.

Further, it is not clear that Streri and Gentaz's experiment shows that the ability to transfer shape information across touch and sight is in fact innate. Rather, it could be that the newborns have learned to correlate distinct visual and tactile sensations of the same shapes. Cylinders and prisms are saliently different; cylinders have curved surfaces, whereas prisms do not, and prisms have vertices, whereas cylinders do not. So, if infants can correlate tactile and visual sensations of curved surfaces and tactile and visual sensations of vertices, that is sufficient for them to perform Streri and Gentaz's task.

There is no reason to think these infants have had insufficient exposure to curved surfaces and vertices for them to learn to correlate visual and tactile sensations of curved surfaces and vertices, respectively. Infants both see and feel many, e.g., curved surfaces, such as those of their own bodies, and those of their mothers', doctors', and nurses' bodies. So there is no reason to think that the ability to transfer information about cylinders and prisms from touch to sight is innate.

In fact, the newborn infants could arguably perform Streri and Gentaz's task without correlating both visual sensations of curved surfaces with tactile sensations of curved surfaces and visual sensations of vertices with tactile sensations of vertices. If an infant has learned to correlate visual and tactile

sensations of curved surfaces only, that infant could perform the crossmodal-transfer task. If the infant is tactually habituated to a cylinder, and is then shown both a cylinder and a prisms, the infant will look longer and more frequently at the prism, since it is novel. On the other hand, an infant who has been tactually habituated to a prism might look longer at the cylinder simply because that infant just had a tactile sensation that is different from the tactile sensation the infant associates with visual sensations of a cylinder; in this case, that the infant looks longer and more frequently at the cylinder reflects this discrepancy, it does not reflect transfer of information about the prism from touch to sight. So an infant need not learn much to perform this task; he or she need only learn to correlate visual and tactile sensations of one feature, e.g., a curved surface. And it is unclear why one would think infants cannot do so in their first day of life. Without ruling out this explanation, Streri and Gentaz have not shown that newborns have an innate ability to transfer shape information from touch to sight.

#### 10. Crossmodal Shape Recognition and Modality Specificity

Further, other experiments on infants' abilities to recognize shapes across sensory modalities provide support for the view that the mental qualities of visual and tactile sensations pertaining to shape are different. Streri and Marie-Germaine Pêcheux (1986) found that 5-month-old infants tactually recognize shapes they have seen but not felt before, but they cannot visually recognize shapes they have felt but not seen. And Streri (1987) found that 2-3-month-old

infants visually recognize shapes they have felt but not seen, but cannot tactually recognize shapes they have seen but not felt. These asymmetries in crossmodal shape recognition, I'll argue, strongly support the view that mental qualities pertaining to shape are modality specific.

To test whether infants visually recognize shapes they have felt but not seen, Streri and Streri and Pêcheux used the same paradigm as in the experiments described above. Infants first manually explored a single shape, either a cylinder or a prism, that they were prevented from seeing. Then, once the tactile stimulus was removed, the infants were shown both a prism and a cylinder, and the experimenters monitored where the infants looked and how long they looked there.

To test whether infants tactually recognize shapes they have seen but not felt, they are first shown one of two shapes, e.g., a cylinder or a prism, that they are prevented from touching. After visual habituation, the visual stimulus is removed, and infants manually explore either a novel-shaped object or an object of the shape they were visually habituated to. The experimenter records how long the infant manually explores the object. Longer manual exploration of novel shapes indicates tactile recognition of the shape infants were visually habituated to.

Using these methods, Streri and Pêcheux (1986) found 5-month-old infants tactually recognize shapes they have seen but not felt, but they do not visually recognize shapes they have felt but not seen. And Streri (1987) found

that 2-3-month-old infants visually recognize shapes they have felt but not seen, but they do not tactually recognize shapes they have seen but not felt. However, they found that both 5-month-olds and 2-3-month-olds in control groups visually recognize shapes that they have previously seen and tactually recognize shapes that they have previously felt. So we cannot explain these asymmetries in crossmodal shape recognition in terms of infants' failure to see or feel shapes, or to remember shapes they have seen or felt.

These asymmetries in infants' crossmodal shape recognition arguably show that the mental qualities of visual and tactile sensations pertaining to shape are distinct. If such mental qualities were the same in sight and touch, crossmodal shape recognition would be automatic; it would make no difference to shape recognition which, a visual or tactile sensation of the shape, one had earlier. But these experiments show that one's ability to recognize a shape one has encountered before does in fact depend on which modality one previously sensed that shape, at least during certain developmental stages.

One might argue that, even if shapes\* were amodal, it could be that crossmodal shape transfer would not be automatic. Rather, it could be that to perform crossmodal shape transfer one must abstract from collateral differences between visual and tactile sensations of the same shapes. Visual sensations of shape have colors\*, which tactile sensations of shape do not have. And tactile sensations of shape have temperatures\*, textures\*, and pressures\*, which visual sensations of shape do not have. So it could be that to coordinate visual and

tactile sensations of the same shape, one must abstract from such modality-specific mental qualities to perform crossmodal shape recognition.

But it is unclear why shapes\* couldn't play their roles in perception without one's abstracting from other mental qualities. Also, the view that one must abstract from such modality-specific mental qualities is compatible with both the view that shapes\* are amodal and the view that they are modality specific. But the view that one must abstract from those modality-specific mental qualities suggests that shapes\* are intricately related to those modality-specific mental qualities. The advocate of the view that shapes\* are amodal must account for that intricate relationship between shapes\* and modality-specific mental qualities while also taking into account the need to abstract away from the modality-specific mental qualities for shapes\* to enable crossmodal shape transfer. It isn't clear how one would do so. In the previous chapter, I argued that the intricate relationship between shapes\* and modality-specific mental qualities such as colors\* and textures\* is best explained in terms of the view that shapes\* are determined by the boundaries of such modality-specific mental qualities. If so, shapes\* are modality specific, not amodal.

Further, it is not clear how we could explain the asymmetries in crossmodal shape transfer on the view that one must abstract from such modality-specific mental qualities to perform such transfer. One might argue that such a view could help explain failures of crossmodal shape transfer in terms of one's failure to abstract from modality-specific mental qualities. If one cannot

abstract from, e.g., the textures\* of one's sensation of a smooth cylinder, then perhaps one will not visually recognize the cylinder when one subsequently sees it. And it could be that infants must learn to perform such abstractions. But presumably one would also have to abstract from the textures\* of a tactile sensation of a cylinder to recognize that shape in a visual-to-tactile trial. And the infants who fail to perform tactile-to-visual shape transfer are able to perform visual-to-tactile shape transfer. So the asymmetry in those infants' crossmodal shape transfer abilities are not due to such failures of abstraction. The advocate of the view that shapes\* are amodal must therefore explain why one can fail to abstract from modality-specific mental qualities in, e.g., tactile-to-visual trials but not in visual-to-tactile trials. It is unclear how such an account would go.

However, if mental qualities representing shape are different in sight and touch, visual and tactile sensations of the same shape must be coordinated to enable crossmodal shape recognition. And it could be that whatever mechanism coordinates visual and tactile sensations of shape sometimes operates asymmetrically, e.g., correlating prior visual sensations with subsequent tactile sensations, but not prior tactile sensations with subsequent visual sensations. Perhaps separate processes underlie visual-to-tactile shape recognition and tactile-to-visual shape recognition, and one process can be suppressed while the other remains active, resulting in the asymmetrical crossmodal shape recognition of 2-3-month-olds and 5-month-olds. For example, it could be that there are two separate processing pathways, one of which enables the transfer of shape

information from visual processing centers to tactile processing centers, but not from tactile to visual processing centers, and another pathway leading from tactile processing centers to visual processing centers. If so, one of those pathways could be rendered inoperative while the other is left operative, resulting in an asymmetry in crossmodal shape transfer.

So, whereas we can account for the asymmetries in infants' crossmodal shape transfer on the view that modality-specific shapes\* must be correlated to enable crossmodal shape transfer, the shapes\* of visual and tactile sensations need not be correlated if they are the same. Further, the asymmetries in infants' crossmodal shape transfer abilities are not due to infants' failure to see or feel shapes, or to remember shapes they previously saw or felt. And we cannot account for the asymmetries on the view that, although shapes\* are amodal, some mechanism must abstract from modality-specific mental qualities in order to enable crossmodal shape transfer. The best explanation of these asymmetries is thus that shapes\* are modality specific, not amodal.

#### 11. Neural Tactile and Visual Representations of Shape

Thomas James, Karin Harman James, Keith Humphrey, and Melvyn Goodale (2006) argue that recent neurophysiological experiments show that visual and tactile representations of shape are the same. Perhaps, one might argue, these experiments support the view that the mental qualities of visual and tactile sensations pertaining to shape are the same. If visual and tactile shape

perception involve exactly the same representations of shape, the properties of those representations are the same.

I'll argue that these experiments fail to show that visual and tactile sensations of the same shapes have amodal mental qualities in common, in virtue of which they are sensations of those same shapes.

Experiments using functional magnetic resonance imaging (fMRI) to monitor neural activity show that tactile recognition of shape activates areas of visual extrastriate cortex, whereas no such activation was found with subjects in control groups in which no shape recognition occurred (Amedi, Jacobson, Hendler, Malach, and Zahary, 2002; Amedi, Malach, Hendler, Peled, and Zohary, 2001; Deibert, Kraut, Kremen, and Hart, 1999). And other experiments show that applying transcranial magnetic stimulation (TMS) to those same extrastriate visual areas suppresses one's ability to tactually identify the orientations of gratings (Zangaladze, Epstein, Grafton, and Sathian, 1999). Since TMS applied to the occipital cortex disrupts one's ability to feel the orientations of stimuli, this suggests that that area of visual cortex is necessary for tactile spatial awareness.

James, Keith Humphrey, Joseph Gati, Philip Servos, Ravi Menon, and Goodale (2002) tested whether, during a visual shape recognition task, prior tactile exposure to shapes and prior visual exposure to shapes result in equivalent increases in neural activation in those areas of visual extrastriate cortex shown to be active during both visual and tactile shape perception. James et al. hypothesized that if prior tactile exploration of a shape and prior visual

exploration of a shape do in fact cause equivalent increases in the activity of lateral occipital cortex (LOC), then the effects of earlier tactile and visual processing on such activation is equivalent. This result, they argue, would challenge the view that tactile representations of shape occur outside of LOC and only indirectly effect activation in LOC. If tactile representations indirectly affect LOC, they argue, tactile shape priming would not activate LOC to the same extent as visual priming, since whatever processing mediated tactile shape processing and the processing in LOC would arguably make tactile priming of LOC less efficient than visual priming. So, they hypothesized, if visual and tactile shape primes cause equivalent activation of LOC during visual shape recognition, tactile representations of shape do not indirectly activate LOC. Rather, they reason, if tactile and visual primes have an equivalent effect on LOC, tactile and visual representations of shape are the same.

Before subjects were scanned by fMRI, James et al. presented them with 16 tactile stimuli and 16 visual stimuli that differed in shape from the tactile stimuli. After this priming stage, subjects were presented with visual images of these 32 objects along with 16 novel objects while being scanned by an fMRI. Subjects were to look at the images, but to do nothing else; i.e., they were instructed to refrain from reacting to the images in any way. The experimenters measured the priming effects of the previously felt and previously seen shapes by monitoring the activation levels those shapes produced in LOC compared to the activation levels caused by novel shapes.

James et al. found that tactile shape priming and visual shape priming produced significant and equivalent levels of activation of LOC. Shapes the subjects had seen and those they had felt during the priming stage produced equivalent levels of activation in LOC, and they produced higher levels of activation than the 16 novel objects did. James et al. take these results to show that vision and touch exploit the same representations of shape in LOC, and that neither involves other representations occurring prior to the activation of LOC. If so, the properties of visual and tactile representations pertaining to shape are the same, so the mental qualities of visual and tactile sensations pertaining to shape are the same.

But it could be that both tactile and visual processing indirectly activate LOC. James et al. assume that no visual representations of shape occur before activation of LOC presumably because LOC is located in what is widely held to be visual cortex. But there is visual processing that feeds forward into LOC, i.e., in both primary visual cortex, V1, and in the lateral geniculate nucleus, LGN, a subcortical area that feeds forward into V1. Perhaps that early visual processing is equivalent to the tactile shape processing feeding into LOC. If so, it could be that both visual and tactile representations of shape occur prior to the activation of LOC. And those visual and tactile representations of shape could differ from each other but still have equivalent effects on the activation of LOC during visual shape recognition.

Also, in addition to feedforward projections from V1 to extrastriate areas, there are feedback projections from extrastriate areas to V1 (Lamme, 2004; Tong, 2003; Bullier, 2001; Pascual-Leone & Walsh, 2001; Hupé et al. 1998; Cowey & Walsh, 2000). Perhaps the visual activation of LOC detected by the fMRI is the product of a significant amount of recurrent processing occurring between V1 and extrastriate cortex. It could be that the amount of such recurrent processing is equivalent to the tactile shape processing that primes activation of LOC, in which case distinct visual and tactile shape processing could activate LOC to the same extent during a subsequent visual shape recognition task.

Further, it could be that tactile and visual shape priming do in fact cause different levels of activation in LOC during visual shape recognition, but that difference is below the threshold of fMRI. If so, fMRI would fail to detect the difference in activation levels that tactile and visual shape processing cause in LOC.

So James et al. fail to show that visual and tactile representations of shape are the same. And they do not show that the mental qualities of visual and tactile sensations pertaining to shape are the same.

None of the neuroscientific or psychological experiments I have discussed support the view that the mental qualities representing shapes are the same in sight and touch. However, the psychological experiments revealing asymmetries in infants' abilities to transfer shape information across sight and touch support the view that such mental qualities are modality specific. Those experiments

suggest that crossmodal shape transfer rests on the correlation of distinct, modality-specific mental qualities representing shapes.

### **Bibliography**

- Amedi, Amir, Gilad Jacobson, Talma Hendler, Rafael Malach, and Ehud Zohary, "Convergence of Visual and Tactile Shape Processing in the Human Lateral Occipital Complex," Cerebral Cortex 12 (2002): 1202-1212.
- Amedi, Amir, Rafael Malach, Talma Hendler, Sharon Peled, and Ehud Zohary, "Visuo-Haptic Object-Related Activation in the Ventral Visual Pathway," Nature Neuroscience 4 (2001): 324-330.
- Anisfeld, Moshe, "Neonatal Imitation: Review," Developmental Review 11 (1991): 60-97.
- Anisfeld, Moshe, "Only Tongue Protrusion Modeling is Matched by Neonates," Developmental Review 16 (1996): 149-161.
- Anisfeld, Moshe, Gerald Turkewitz, Susan A. Rose, Faigi R. Rosenberg, Faith J. Sheiber, Deborah A. Couturier-Fagan, Joseph S. Ger, and Iris Sommer, "No Compelling Evidence that Newborns Imitate Oral Gestures," Infancy 2, 1 (2001): 111-122.
- Aristotle, de Anima, transl. R. D. Hicks, Cambridge: Cambridge University Press, 1907.
- Armstrong, David M., A Materialist Theory of the Mind, New York: Humanities Press, 1968.
- Bach-y-Rita, Paul, Brain Mechanisms in Sensory Substitution, New York: Academic Press, 1972.
- Baylis, Gordon and Jon Driver, "Visual Attention and Objects: Two-object Cost with Equal Convexity," Journal of Experimental Psychology: Human Perception and Performance 19 (1993): 451-470.
- Beck, Diane, Geraint Rees, Christopher Frith, and Nilli Lavie, "Neural Correlates of Change Detection and Change Blindness," Nature Neuroscience 4, 6 (June 2001): 645-650.
- Bermúdez, José Luis, The Paradox of Self-Consciousness, Cambridge: MIT Press, 1998.
- Berkeley, George, "An Essay Towards a New Theory of Vision," edited from the fourth edition (1732) by M. R. Ayers, George Berkeley: Philosophical Works, Rutland, VT: Charles E. Tuttle Co., Inc., 1975, pp. 3-59.

- Bertelson, Paul, Francesco Pavani, Elisabetta Ladavas, Jean Vroomen, and Béatrice de Gelder, "Ventriloquism in Patients With Unilateral Visual Neglect," Neuropsychologia 38, 12 (October 2000): 1634-1642.
- Blackmore, Susan J., Gavin Brelstaff, Kay Nelson, and Tom Troscianko, "Is the Richness of Our Visual World an Illusion? – Trans-saccadic Memory for Complex Scenes," Perception 24 (1995): 1075-1081c.
- Block, Ned, "Review of Alva Noë, Action in Perception," Journal of Philosophy 102, 5 (May 2005): 259-272.
- Breitmeyer, Bruno, Haluk Ogmen, and Jian Chen, "Unconscious Priming by Color and Form: Different Processes and Levels," Consciousness and Cognition 13, 1 (March 2004): 138-157.
- Brewer, Bill, Perception and Reason, New York: Oxford University Press, 1999.
- Brown, Kathleen W. and Allen W. Gottfried, "Cross-Modal Transfer of Shape in Early Infancy: Is there Reliable Evidence?" In eds. L. P. Lipsitt and R. Rovee-Collier, Advances in Infancy Research, Norwood, NJ: Ablex, 1986, pp. 163-170.
- Bullier, Jean, "Feedback Connections and Conscious Vision," Trends in Cognitive Sciences 5, 9 (September 2001): 369-370.
- Campbell, John, "Molyneux's Problem," in Perception: Philosophical Issues, 7, 1996, ed. Enrique Villanueva, Atascadero, California: Ridgeview Publishing Company, 1996a, pp. 301-318.
- Campbell, John, "Shape Properties, Experience of Shape and Shape Concepts," in Perception: Philosophical Issues, 7, 1996, ed. Enrique Villanueva, Atascadero, California: Ridgeview Publishing Company, 1996b, pp. 351-363.
- Campbell, John, Reference and Consciousness, Oxford: Oxford University Press, 2002.
- Campbell, John, "Information-Processing, Phenomenal Consciousness and Molyneux's Question," in ed. José Luis Bermúdez, Thought, Reference and Experience: Themes from the Philosophy of Gareth Evans, Oxford: Oxford University Press, 2005, pp. 195-219.

- Chalmers, David, The Conscious Mind, New York: Oxford University Press, 1996.
- Clark, Andy, "Visual Experience and Motor Action: Are the Bonds Too Tight?" Philosophical Review 110 (2001): 495-519.
- Clark, Austen, Sensory Qualities, Oxford: Oxford University Press, 1993.
- Clark, Austen, "Three Varieties of Visual Field," Philosophical Psychology 9, 4 (1996): 477-495.
- Clark, Austen, A Theory of Sentience, Oxford: Oxford University Press, 2000.
- Clark, Austen, "Feature-Placing and Proto-Objects," Philosophical Psychology 17, 4 (December 2004): 443-469.
- Cohen, Asher and Richard Ivry, "Illusory Conjunctions Inside and Outside the Focus of Attention," Journal of Experimental Psychology: Human Perception and Performance 15, 4 (November 1989): 650-663.
- Cohen, Asher and Robert D. Rafal, "Attention and Feature Integration: Illusory Conjunctions in a Patient With a Parietal Lobe Lesion," Psychological Science 2, 2 (March 1991): 106-110.
- Cohen, Jonathan, "Objects, Places, and Perception," Philosophical Psychology 17, 4 (December 2004): 471-495.
- Cowey, Alan and V. Walsh, "Magnetically Induced Phosphenes in Sighted, Blind and Blindsighted Observers," Neuroreport: For Rapid Communication of Neuroscience Research 11, 14 (September 2000): 3269-3273.
- Crane, Tim, "The Waterfall Illusion," Analysis 48 (1988a): 142-147.
- Crane, Tim, "Concepts in Perception," Analysis 48 (1988b): 150-153.
- Cussins, Adrian, "The Connectionist Construction of Concepts," in The Philosophy of Artificial Intelligence, ed. M. Boden, Oxford: Oxford University Press, 1990.
- Deibert, Ellen, Michael Kraut, Sarah Kremen, and John Hart, "Neural Pathways in Tactile Object Recognition," Neurology 52 (1999): 1413-1417.
- Dennett, Daniel C., "Two Approaches to Mental Images," in Brainstorms, Cambridge: MIT Press, 1978.

- Dennett, Daniel C., Consciousness Explained, Boston: Little, Brown, 1991.
- Dennett, Daniel C., Sweet Dreams, Cambridge: MIT Press, 2005.
- Dretske, Fred, Naturalizing the Mind, Cambridge: MIT Press, 1995.
- Dretske, Fred, "Change Blindness," Philosophical Studies 120 (2004): 1-18.
- Driver, Jon, Greg Davis, Charlotte Russell, Massimo Turatto, and Elliot Freeman, "Segmentation, Attention and Phenomenal Visual Objects," Cognition 80, 1-2 (June 2001): 61-95.
- Egley, Robert, Jon Driver, and Robert Rafal, "Shifting Visual Attention Between Objects and Locations: Evidence from Normal and Parietal Lesion Subjects," Journal of Experimental Psychology: General 123, 2 (June 1994): 161-177.
- Ericksen, B. A. and C. W. Ericksen, "Effects of Noise Letters upon the Identification of a Target Letter in a Nonsearch Task," Perception and Psychophysics 16 (1974): 143-149.
- Evans, Gareth, The Varieties of Reference, Oxford: Oxford University Press, 1982.
- Fernandez-Duque, Diego, Giordana Grossi, Ian Thornton, and Helen Neville, "Representation of Change: Separate Electrophysiological Markers of Attention, Awareness, and Implicit Processing," Journal of Cognitive Neuroscience 15, 4 (2003): 491-507.
- Fernandez-Duque, Diego and Ian M. Thornton, "Change Detection Without Awareness: Do Explicit Reports Underestimate the Representation of Change in the Visual System?" Visual Cognition 7 (2000): 324-344.
- Fernandez-Duque, Diego and Ian Thornton, "Explicit Mechanisms Do Not Account for Implicit Localization and Identification of Change: An Empirical Reply to Mitroff et al. (2002)," Journal of Experimental Psychology: Human Perception and Performance 29, 5 (2003): 846-858.
- Fodor, Jerry A., The Modularity of Mind, Cambridge: MIT Press, 1983.
- Fodor, Jerry A. and Zenon W. Pylyshyn, "How Direct is Visual Perception?: Some Reflections on Gibson's 'Ecological Approach'," Cognition 9 (1981): 139-196.

- Friedman-Hill, Stacia R., Lynn C. Robertson, Anne Treisman, "Parietal Contributions to Visual Feature Binding: Evidence from a Patient With Bilateral Lesions," Science 269, 5225 (August 1995): 853-855.
- Gibson, James J., The Senses Considered as Perceptual Systems, Boston: Houghton Mifflin Company, 1966.
- Gibson, James J., The Ecological Approach to Visual Perception, Boston: Houghton-Mifflin Company, 1979.
- Goodale, Melvyn A. and Kelly Murphy, "Action and Perception in the Visual Periphery," in P. Theier and H.-O. Karnath, eds., Parietal Lobe Contributions to Orientation in 3 D Space, New York: Springer, 1997, pp. 447-461.
- Goodman, Nelson, The Structure of Appearance, third edition, Dordrecht: Reidel, 1977.
- Grimes, John, "On the Failure to Detect Changes in Scenes Across Saccades," in Perception, ed. Kathleen Akins, New York: Oxford University Press, 1996.
- Haffenden, Angela M., Karen C. Schiff, and Melvyn A. Goodale, "The Dissociation Between Perception and Action in the Ebbinghaus Illusion: Nonillusory Effects of Pictorial Cues on Grasp," Current Biology 11 (February 2001): 177-181.
- Haggard, Patrick, "Conscious Intention and Motor Cognition," Trends in Cognitive Sciences 9, 6 (June 2005): 290-295.
- Helmholtz, H. von., Treatise on Physiological Optics, vol. 3, Translated from German by J. P. C. Southall, New York: Dover, 1867/1962.
- Hollingworth, Andrew, Carrick C. Williams, & John M. Henderson, "To See and Remember: Visually Specific Information Is Retained In Memory From Previously Attended Objects In Natural Scenes," Psychonomic Bulletin & Review 8, 4 (2001): 761-768.
- Houck, Michael R. and James E. Hoffman, "Conjunction of Color and Form Without Attention: Evidence From an Orientation-Contingent Color Aftereffect," Journal of Experimental Psychology: Human Perception and Performance 12, 2 (May 1986): 186-199.

- Humphrey, Nicholas, Consciousness Regained: Chapters In the Development of Mind, Oxford: Oxford University Press, 1983.
- Hupé, J. M., A. C. James, B. R. Payne, S. G. Lomber, P. Girard, and J. Bullier, "Cortical Feedback Improves Discrimination Between Figure and Background by V1, V2 and V3 Neurons," Nature 394 (August 1998): 784-787.
- Hurley, Susan, Consciousness in Action, Cambridge: Harvard University Press, 1998.
- Jackson, Frank, Perception: A Representative Theory, Cambridge: Cambridge University Press, 1977.
- James, Thomas, Karin Harman James, Keith Humphrey, and Melvyn A. Goodale, "Do Visual and Tactile Object Representations Share the Same Neural Substrate?" in eds. Morton A. Heller and Soledad Ballesteros, "Touch and Blindness: Psychology and Neuroscience," Mahwah, N.J.: Lawrence Erlbaum, 2006.
- James, Thomas, Keith Humphrey, Joseph Gati, Philip Servos, Ravi Menon, and Melvyn A. Goodale, "Haptic Study of Three-Dimensional Objects Activates Extrastriate Visual Areas," Neuropsychologia 40 (2002): 1706-1714.
- Jones, Barry T., Ben C. Jones, Helena Smith, and Nicola Copley, "A Flicker Paradigm for Inducing Change Blindness Reveals Alcohol and Cannabis Information Processing Biases in Social Users," Addiction 98 (2003): 235-244.
- Jones, Susan, "Imitation or Exploration? Young Infants' Matching of Adults' Oral Gestures," Child Development 67 (1996): 1952-1969.
- Kahneman, Daniel, Anne Treisman, and Brian J. Gibbs, "The Reviewing of Object Files: Object-Specific Integration of Information," Cognitive Psychology 24, 2 (April 2002): 175-219.
- Kentridge, Robert W., Charles A. Heywood, and Lawrence Weiskrantz, "Spatial Attention Speeds Discrimination Without Awareness in Blindsight," Neuropsychologia 42, 6 (2004): 831-835.
- Kosslyn, Stephen, Image and Brain, Cambridge: MIT Press, 1994.

- Krikpe, Saul, Naming and Necessity, Cambridge: Harvard University Press, 1980.
- Kripke, Saul, Wittgenstein on Rules and Private Language, Oxford: Blackwell Publishing, 1982.
- Laloyaux, Cédric, Arnaud Destrebecqz, and Axel Cleeremans, "Implicit Change Identification: A Replication of Fernandez-Duque and Thornton (2003)," Journal of Experimental Psychology: Human Perception and Performance, forthcoming.
- Laloyaux, Cédric, Christel Devue, Elodie David, and Axel Cleeremans, "Change Blindness to Gradual Changes in Facial Expressions," submitted.
- Lamme, Victor A. F., "Separate Neural Definitions of Visual Consciousness and Visual Attention; a Case for Phenomenal Awareness," Neural Networks 17, 5-6 (June and July 2004): 861-872.
- Levin, Daniel T. and Daniel J. Simons, "Failure to Detect Changes to Attended Objects in Motion Pictures," Psychonomic Bulletin & Review 4 (1997): 501-6.
- Levine, Joseph, "Materialism and Qualia: The Explanatory Gap," Pacific Philosophical Quarterly 64 (1983): 354-361.
- Levine, Joseph, Purple Haze, New York: Oxford University Press, 2001.
- Levine, Michael W., Levine and Shefner's Fundamentals of Sensation and Perception, 3rd ed., Oxford: Oxford University Press, 2000.
- Loar, Brian, "Comments on John Campbell, 'Molyneux's Question'," in Perception: Philosophical Issues, 7, 1996, ed. Enrique Villanueva, Atascadero, California: Ridgeview Publishing Company, 1996a, pp. 301-318.
- Locke, John, An Essay Concerning Human Understanding, edited from the fourth (1700) edition by Peter H. Nidditch, Oxford: Oxford University Press, 1975.
- Lu, Chen-Hui and Robert W. Proctor, "The Influence of Irrelevant Location Information on Performance: A Review of the Simon and Spatial Stroop Effects," Psychonomic Bulletin and Review 2 (1995): 174-207.
- Lycan, William G., Consciousness and Experience, Cambridge: MIT, 1996.

- Mach, Ernst, Analysis of Sensation, New York: Dover, 1906/1959.
- Mack, Arien, "Is the Visual World a Grand Illusion? A Response," Journal of Consciousness Studies, Vol. 9, No. 5-6, (2002): 102-110.
- Mack, Arien and Irvin Rock, Inattentional Blindness, Cambridge: MIT Press, 1998.
- Marcel, Anthony J., "Conscious and Unconscious Perceptions: An Approach to the Relations between Phenomenal Experience and Perceptual Processes," Cognitive Psychology 15 (1983): 238-300.
- Marr, David, Vision, New York: W. H. Freeman, 1982.
- Matthen, Mohan, "Features, Places, and Things: Reflections on Austen Clark's Theory of Sentience," Philosophical Psychology 17, 4 (December 2004): 497-518.
- Maurer, Daphne and Catherine Mondloch, "Synesthesia: A Stage of Normal Infancy?" in ed. S. C. Masin, Fechner Day 96. Proceedings of the Twelfth Annual Meeting of the International Society for Psychophysics, Padua, Italy: The International Society for Psychophysics, 1996, pp. 107-112.
- Maurer, Daphne, Christine L. Stager, and Catherine J. Mondloch, "Cross-Modal Transfer of Shape is Difficult to Demonstrate in One-Month-Olds," Child Development 70, 5 (September/October 1999): 1047-1057.
- McBeath, Michael K., Dennis M. Shaffer, Mary K. Kaiser, "How Baseball Outfielders Determine Where to Run to Catch Fly Balls," Science 268, 5210 (April 1995): 569-573.
- McLeod, Peter, Nick Reed, and Zoltan Dienes, "The Optic Trajectory Is Not a Lot of Use if You Want to Catch the Ball," Journal of Experimental Psychology: Human Perception and Performance 28, 6 (December 2002): 1499-1501.
- McConkie, George and David Zola, "Is Visual Information Integrated Across Successive Fixations in Reading?" Perception and Psychophysics 25 (1979): 221-224.

- McConkie, George, David Zola, G. S. Wolverton, and D. D. Burns, "Eye Movement Contingent Display Control in Studying Reading," Behavior Research Methods and Instrumentation 4 (1978): 529-544.
- McDowell, John, Mind and World, Cambridge: Harvard University Press, 1994.
- Meehan, Douglas B., "Spatial Experience, Sensory Qualities, and the Visual Field," in eds. Johanna D. Moore and Keith Stenning, Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society, Mahwah, NJ: Lawrence Erlbaum Associates, 2001, pp. 623-627.
- Meehan, Douglas B., "Qualitative Character and Sensory Representation," Consciousness and Cognition, 11, 4 (December 2002): 630-641.
- Meehan, Douglas B., "Phenomenal Space and the Unity of Conscious Experience," Psyche, 9, 12 (May 2003), at <http://psyche.cs.monash.edu.au/symposia/dainton/meehan.html>
- Meltzoff, Andrew, "Molyneux's Babies: Cross-modal Perception, Imitation and the Mind of the Preverbal Infant," in eds. Naomi Eilan, Rosaleen McCarthy, and Bill Brewer, Spatial Representation, Oxford: Oxford University Press, 1993, pp. 219-235.
- Meltzoff, Andrew and M. Keith Moore, "Imitation of Facial and Manual Gestures by Human Neonates," Science 198 (1977): 75-78.
- Meltzoff, Andrew and M. Keith Moore, "Newborn Infants Imitate Adult Facial Gestures," Child Development 54 (1983): 702-709.
- Meltzoff, Andrew and Richard Borton, "Intermodal Matching by Human Neonates," Nature 282 (1979): 403-404.
- Milner, A. David and Melvyn A. Goodale, The Visual Brain in Action, New York: Oxford University Press, 1995.
- Mitroff, Stephen R., Daniel J. Simons, and Steven L. Franconeri, "The Siren Song of Implicit Change Detection," Journal of Experimental Psychology: Human Perception and Performance 28, 4 (2002): 798-815.
- Mitroff, Stephen, Daniel Simons, and Daniel Levin, "Nothing Compares 2 Views: Change Blindness Can Occur Despite Preserved Access to the Changed Information," Perception and Psychophysics 66, 8 (2004): 1268-1281.

- Nissen, Mary Jo, "Accessing Features and Objects: Is Location Special?" in eds. M. I. Posner and O. S. Marin, Attention and Performance XI, Hillsdale, N. J.: Erlbaum, pp. 205-219.
- Noë, Alva, Action in Perception, Cambridge: MIT Press, 2004.
- Noë, Alva, "What Does Change Blindness Teach Us about Consciousness?" Trends in Cognitive Sciences 9, 5 (May 2005): p. 218.
- Ogawa, Hirokazu and Akihiro Yagi, "The Implicit Processing in Multiple Object Tracking," Technical Report on Attention and Cognition 1, 10 (2002).
- O'Regan, J. Kevin, "Solving the 'Real' Mysteries of Visual Perception: The World as an Outside Memory," Canadian Journal of Psychology 46, 3 (1992): 461-488.
- O'Regan, J. Kevin, Heiner Deubel, James J. Clark, and Ronald A. Rensink, "Picture Changes During Blinks: Looking Without Seeing and Seeing Without Looking," Visual Cognition 7, 1-3 (2000): 191-211.
- O'Regan, Kevin and Alva Noë, "A Sensorimotor Account of Vision and Visual Consciousness," Behavioral and Brain Sciences 24, 5 (2001): 939-973.
- O'Regan, Kevin, Ronald A. Rensink, & J. J. Clark, "'Mud Splashes' Render Picture Changes Invisible," Investigative Ophthalmology and Visual Science 37 (1996): S213.
- Pascual-Leone, Alvaro and Vincent Walsh, "Fast Backprojections From the Motion to the Primary Visual Area Necessary for Visual Awareness," Science 292, 5516 (April, 2001): 510-512.
- Peacocke, Christopher, Sense and Content, Oxford: Oxford University Press, 1983.
- Peacocke, Christopher, A Study of Concepts, Cambridge: MIT Press, 1992.
- Peacocke, Christopher, "Does Perception Have a Nonconceptual Content?" The Journal of Philosophy 98, 5 (May 2001): 239-264.
- Pêcheux, Marie-Germaine, J-C Lepecq, and P. Salzarulo, "Oral Activity and Exploration in 1- to 2-month-old Infants," British Journal of Developmental Psychology 6 (1988): 245-256.

- Perenin, Marie-Thérèse and A. Vighetto, "Optic Ataxia: A Specific Disorder in Visuomotor Coordination," in eds. A. Hein and M. Jeannerod, Spatially Oriented Behavior, New York: Springer-Verlag, 1988, pp. 305-326.
- Perry, John, "The Problem of the Essential Indexical," Nous 13, 1 (March 1979): 3-21.
- Pessoa, Luiz and Leslie Ungerleider, "Neural Correlates of Change Detection and Change Blindness in a Working Memory Task," Cerebral Cortex 14 (2004): 511-520.
- Pitcher, George, A Theory of Perception, Princeton, NJ: Princeton University Press, 1971.
- Posner, Michael, Chronometric Explorations of Mind, Hillsdale, N. J.: Erlbaum, 1978.
- Posner, Michael, "Orienting of Attention," Quarterly Journal of Experimental Psychology 32, 1 (February 1980): 3-25.
- Prinz, Jesse, "A Neurofunctional Theory of Consciousness," in eds. Andrew Andrew and Kathleen Akins, Cognition and the Brain, Cambridge: Cambridge University Press, 2005, pp. 381-398.
- Pylyshyn, Zenon, "Visual Indexes and Nonconceptual Reference," unpublished.
- Pylyshyn, Zenon, Seeing and Visualizing, Cambridge: MIT Press, 2003.
- Pylyshyn, Zenon and Ron Storm, "Tracking Multiple Independent Targets: Evidence for a Parallel Tracking Mechanism," Spatial Vision 3 (1988): 1-19.
- Quine, Willard V. O., "On What There Is," in Willard V. O. Quine, From a Logical Point of View, Cambridge: Harvard University Press, 1953, pp. 1-19.
- Ratcliff, Graham and G. A. B. Davies-Jones, "Defective Visual Localization in Focal Brain Wounds," Brain 95 (1972): 49-60.
- Rensink, Ronald, "Seeing, Sensing, and Scrutinizing," Vision Research 40 (2000): 1469-1487.
- Rensink, Ronald, J. Kevin O'Regan, and James Clark, "To See or Not to See: The Need for Attention to Perceive Changes in Scenes," Psychological Science 8 (1997): 368-373.

- Riddoch, F., "Visual Disorientation in Homonymous Half-Fields," Brain 58 (1935): 376-382.
- Robertson, Lynn, Anne Treisman, Stacia Friedman-Hill, and Marcia Grabowecky, "The Interaction of Spatial and Object Pathways: Evidence from Balint's Syndrome," Journal of Cognitive Neuroscience 9, 3 (May 1997): 295-317.
- Rock, Irvin, Introduction to Perception, New York: Macmillan, 1975.
- Rock, Irvin, Indirect Perception, Cambridge: MIT Press, 1997.
- Rosenthal, David M., "The Independence of Consciousness and Sensory Quality," in Consciousness: Philosophical Issues, 1, 1991, ed. Enrique Villanueva, Atascadero, California: Ridgeview Publishing Company, 1991, pp. 15-36.
- Rosenthal, David M., "A Theory of Consciousness," in The Nature of Consciousness, eds. Ned Block, Owen Flanagan, and Güven Güzeldere, Cambridge: MIT Press, 1997, 729-753.
- Rosenthal, David M., "Sensory Quality and the Relocation Story," Philosophical Topics 26, 1 and 2 (Spring and Fall 1998): 321-350.
- Rosenthal, David M., "The Colors and Shapes of Visual Experiences," in Consciousness and Intentionality: Models and Modalities of Attribution, ed. Denis Fisette. Dordrecht: Kluwer, 1999, 95-118.
- Rosenthal, David M., "Color, Mental Location, and the Visual Field," Consciousness and Cognition 10, 1 (March 2001): 85-93.
- Rosenthal, David M., "Sensory Qualities, Consciousness, and Perception," in David M. Rosenthal, Consciousness and Mind, Oxford: Clarendon Press, 2005.
- Russell, Charlotte and Jon Driver, "New Indirect Measures of 'Inattentive' Visual Grouping in a Change-detection Task," Perception & Psychophysics 67 (2005): 606-623.
- Russell, Bertrand, The Problems of Philosophy, Oxford: Oxford University Press, 1912.
- Scholl, Brian J., "Objects and Attention: The State of the Art," Cognition 80, 1-2 (June 2001): 1-46.

- Scholl, Brian J. and Zenon Pylyshyn, "Tracking Multiple Items Through Occlusion: Clues to Visual Objecthood," Cognitive Psychology 38, 2 (1999): 259-290.
- Scholl, Brian J., Zenon Pylyshyn, and Steven Franconeri, "The Relationship Between Property-Encoding and Object-Based Attention: Evidence from Multiple Object Tracking," unpublished.
- Sellars, Wilfrid, Science and Metaphysics: Variations on Kantian Themes. London: Routledge & Kegan Paul, 1968.
- Sellars, Wilfrid, "Empiricism and the Philosophy of Mind," in Science, Perception and Reality, London: Routledge & Kegan Paul, 1963.
- Shoemaker, Sydney, "Functionalism and Qualia," Philosophical Studies XXVII, 5 (May 1975): 292-315.
- Shoemaker, Sydney, "Properties and Causality," in Identity, Cause and Mind, Cambridge: Cambridge University Press, 1984.
- Shoemaker, Sydney, "The Royce Lectures: Self-knowledge and 'Inner Sense'," in The First-Person Perspective and Other Essays, Cambridge: Cambridge University Press, 1996.
- Simon, Richard and A. M. Small, Jr., "Processing Auditory Information: Interference from an Auditory Cue," Journal of Applied Psychology 53 (1969): 433-435.
- Simons, Dan, Steven Franconeri, and Rebecca Reimer, "Change Blindness in the Absence of a Visual Disruption," Perception 29 (2000): 1143-1154.
- Simons, Daniel and Daniel Levin, "Failure to Detect Changes to People in a Real-World Interaction," Psychonomic Bulletin and Review 5 (1998): 644-649.
- Smilek, Daniel, John Eastwood, and Philip Merikle, "Does Unattended Information Facilitate Change Detection?" Journal of Experimental Psychology: Human Perception and Performance 26, 2 (2000): 480-487.
- Strawson, Galen, Mental Reality, Cambridge: MIT Press, 1994.
- Strawson, Peter F., "Kant's Theory of Geometry," in The Bounds of Sense, London: Methuen, 1966.

- Streri, Arlette, "Tactile Discrimination of Shape and Intermodal Transfer in 2- to 3-month-old Infants," British Journal of Developmental Psychology, 5 (1987): 213-220.
- Streri, Arlette and Edouard Gentaz, "Cross-Modal Recognition of Shape from Hand to Eyes in Human Newborns," Somatosensory & Motor Research 20, 1 (2003): 13-18.
- Streri, Arlette and Edouard Gentaz, "Cross-Modal Recognition of Shape from Hand to Eyes and Handedness in Human Newborns," Neuropsychologia 42 (2004): 1365-1369.
- Streri, Arlette and Marie-Germaine Pêcheux, "Vision-to-Touch and Touch-to-Vision Transfer of Form in 5-month-old Infants," British Journal of Developmental Psychology, 4 (1986): 161-167.
- Stroop, J. R., "Studies of Interference in Serial Verbal Reactions," Journal of Experimental Psychology 18 (1935): 643-662.
- Thornton, Ian M. and Diego Fernandez-Duque, "An Implicit Measure of Undetected Change," Spatial Vision 14, 1 (2000): 21-44.
- Thornton, Ian M. and Diego Fernandez-Duque, "Converging Evidence for the Detection of Change Without Awareness," in eds. J. Hyönä, D. P. Munoz, W. Heide, and R. Radach, The Brain's Eye: Neurobiological and Clinical Aspects of Oculomotor Research: Progress in Brain Research, vol. 140, 2002, pp. 99-118.
- Tong, Frank, "Primary Visual Cortex and Visual Awareness," Nature Neuroscience 4 (March 2003): 219-229.
- Treisman, Anne, "Features and Objects: The Fourteenth Annual Bartlett Memorial Lecture," Quarterly Journal of Experimental Psychology A, 40 (1988): 201-237.
- Treisman, Anne, "Feature Binding, Attention and Object Perception," in eds. Glyn W. Humphreys, John Duncan, and Anne Treisman, Attention, Space and Action: Studies in Cognitive Neuroscience, Oxford: Oxford University Press, 1999, pp. 91-111.
- Treisman, Anne and Garry Gelade, "A Feature-Integration Theory of Attention," Cognitive Psychology 12, 1 (January 1980): 97-136.

- Treisman, Anne and Hilary Schmidt, "Illusory Conjunctions in the Perception of Objects," Cognitive Psychology 14, 1 (January 1982): 107-141.
- Turatto, Massimo, Angrilli Alessandro, Veronica Mazza, Carolo Umiltà, and Jon Driver, "Looking Without Seeing the Background Change: Electrophysiological Correlates of Change Detection Versus Change Blindness," Cognition 84 (2002): B1-10.
- Tye, Michael, Ten Problems of Consciousness, Cambridge: MIT Press, 1995.
- Tye, Michael, "Perceptual Experience is a Many-Layered Thing," in Perception: Philosophical Issues 7, 1996, ed. Enrique Villanueva, Atascadero, CA: Ridgeview Publishing Company, 1996, pp. 117-126.
- Viswanathan, Lavanya and Ennio Mingolla, "Dynamics of Attention in Depth: Evidence from Multi-Element Tracking," Perception 31, 12 (2002): 1415-1437.
- Weiskrantz, Lawrence, Consciousness Lost and Found: A Neuropsychological Exploration, Oxford: Oxford University Press, 1997.
- Werner, Steffen and Bjorn Thies, "Is 'Change Blindness' Attenuated by Domain-specific Expertise? An Expert-Novices Comparison of Change Detection in Football Images," Visual Cognition 7, 1-3 (2000): 163-174.
- Williams, Pepper and Daniel Simons, "Detecting Changes in Novel, Complex Three-dimensional Objects," Visual Cognition 7 (2000): 297-322.
- Wittgenstein, Ludwig, Philosophical Investigations, New York: MacMillan, 1953.
- Zangaladze, Andro, Charles M., Epstein, Scott T. Grafton, and K. Sathian, "Involvement of Visual Cortex in Tactile Discrimination of Orientation," Nature 401 (1999): 587-590.