OXFORD

PERCEPTUAL EXPERIENCE

Christopher S. Hill

Perceptual Experience

Perceptual Experience

CHRISTOPHER S. HILL

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

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

First Edition published in 2022

Impression: 1

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

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2022935538

ISBN 978–0–19–286776–6

DOI: 10.1093/oso/9780192867766.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

For my friends and teachers, living and dead

Contents

Preface

The title of this volume is accurate: all of the chapters are concerned, to at least some degree, with perceptual experience, and there is no other topic that has a role of comparable size in the discussion. But parts of the book have a narrower scope than the title might lead one to expect. Thus, visual experience figures much more prominently in most of the chapters than the forms of experience that are associated with other perceptual modalities. Moreover, several other parts of the book have a much broader scope than the title might suggest. For example, Chapter 1 contains an extended presentation and defense of representationalism, the view that all mental states, including those with an experiential dimension, constitutively involve representations; and Chapters 5 and 6 propose general accounts of phenomenology, including several forms of phenomenology that are linked to perception but are not themselves perceptual in character. There are reasons for these departures from what is strictly promised by the title, in addition to the intrinsic interest of the topics. My reason for focusing on visual experience has primarily to do with the comparative extent and depth of the vision science literature. There is much more philosophically relevant knowledge about vision than there is, for example, about audition. As for reasons for expanding the scope of the book beyond perceptual experience, there are two of them. One is that it is often difficult to make a fully convincing case for a claim about perceptual experience unless one can show that it generalizes to other forms of experience. And the other reason is that a truly comprehensive picture of perceptual experience requires complementary accounts of its relations to other mental phenomena.

Although I believe there to be strong intuitive motivation for most of the main claims of the book, the principal arguments often appeal to experimental literatures in cognitive science and neuroscience. For the most part, I eschew purely a priori argumentation. This is because my objective in this volume is not to understand the commonsense concept of perceptual experience, but to understand perceptual experience itself, which, as we will see, appears to be a natural kind. Studies of our commonsense worldview are important. We all want very much to understand the commonsense ways in which members of our species represent the world and think about it. In my view, however, the relationship between the principles governing our commonsense concepts and genuine natural kinds is often quite loose and tenuous, and this is particularly true of the principles that figure in the commonsense theory of mental states and processes. Much of folk psychology is admirably on target, but other parts are deeply misleading, creating snares from which philosophers have only begun, after centuries of confusion, to

extricate themselves. I will be developing a case for this skeptical stance as we proceed.

I hope the present volume will be the first in a pair with the overarching title *Windows of the Mind*, but in an uncertain world it is hard to be confident that this larger ambition will be fulfilled. The first window is of course perceptual experience. The second consists of concepts of all kinds, scientific as well as commonsense, and states that have concepts as their building blocks.

Acknowledgments

I owe much to the knowledge and wisdom of others.

First, I have been helped in developing my ideas, to a degree beyond all acknowledging, by four extraordinary books—Ned Block's *The Border between Seeing and Thinking* (Block 2022), Tyler Burge's *Origins of Objectivity* (Burge 2010), Karen Neander's *The Mark of the Mental* (Neander 2017), and Nicholas Shea's *Representation in Cognitive Science* (Shea 2018).

Second, I have been helped to understand many issues, and saved from many errors, by conversations and/or correspondence with Louise Antony, Murat Aydede, Davide Badre, David J. Bennett, Jacob Beck, Ori Beck, Ned Block, Anthony Brueckner, Alex Byrne, Jonathan Cohen, Robert Cummins, Fulvio Domini, James Dreier, Fred Dretske, Frank Durgin, Chaz Firestone, Ivan Fox, Alvin Goldman, E. J. Green, Louis Gularte, Anil Gupta, Riki Heck, Jaegwon Kim, Joseph Levine, Jack Lyons, John Mazlish, Matthew McGrath, Brian McLaughlin, Karen Neander, Casey O'Callaghan, Adam Pautz, Joshua Schechter, Susanna Schellenberg, Thomas Senor, Ayoob Shahmoradi, Nicholas Shea, Amitai Shenav, Sydney Shoemaker, Susanna Siegel, James Van Cleve, Lee Warren, William Warren, Wayne Wu, and Steven Yamamoto.

Third, my editor, Peter Momtchiloff, selected three outstandingly gifted and generous referees. They have saved me from myself at a large number of places, by spotting errors both small and large, and have also made positive suggestions that have enriched the book considerably. I add that Peter has been helpful in other ways as well, not least in providing encouragement.

Fourth, the book has been lucky to have Henry Clarke as its project editor, Bhavani Govindasamy as its project manager, and Dorothy McCarthy as its copy editor.

Fifth, I am indebted for advice about various literatures to David Abel, Justin Broackes, Alan Gilchrist, Matthew Fulkerson, Roma Patel, James Pryor, Robert Sapolsky, Andrea Simmons, and Albert Webson.

Sixth, I am grateful to Beth Greely for preparing images (Figures 2.2, 2.3, 2.4, 5.1, and 8.2) that I could never have produced myself, and to Carl Granrud for supplying Figure 2.1.

Chapter 2 is an expanded version of an earlier paper ("Appearance and Reality," *Philosophical Issues* (2020) 30: 175–91). In addition, I have excerpted short passages from the following publications: "Ow! The Paradox of Pain," in M. Aydede (ed.), *Pain: New Essays on its Nature and the Methodology of it Study* (Cambridge, MA: MIT Press, 2005), 75–98; "Locating Qualia: Are They in the Mind or in the

Body and the World?" in S. Gozzano and C. S. Hill (eds.), *New Perspectives on Type Identity* (Cambridge University Press, 2012), 127–49; "Perceptual Relativity," *Philosophical Topics* (2016) 44(2): 179–200; "Replies to Byrne, McGrath, and McLaughlin," *Philosophical Studies* (2016) 173: 861–72; "Goldman on Knowledge of Mind," in H. Kornblith and B. McLaughlin (eds.), *Goldman and His Critics* (Wiley, 2016), 259–76; "Perceptual Existentialism Sustained," *Erkenntnis* (2019), https://doi.org/10.1007/s10670-019-00160-z; and "Tribute to Karen Neander," co-authored with Carlotta Pavese, *Biological Theory* (2021) https://doi.org/10.1007/s13752-021-00383-w. The excerpts from these articles are reprinted with permission from the publishers.

List of Figures

The crows and cloughs that wing the midway air
Show scarce so gross as beetles. Half way down
Hangs one that gathers samphire, dreadful trade!
Methinks he shows no bigger than his head.
The fishermen that walk upon the beach
Appear like mice, and yon tall anchoring bark
Diminished to her cock—her cock, a buoy
Almost too small for sight.

King Lear IV, vi

Everything that is represented through a sense is to that extent always appearance.

Kant (1998, p. 189)

1
Representationalism

I. Introduction

Intentionality is the property that a mental state possesses if it is directed on something other than itself. (The term comes from the Latin verb *intendere*, which means *to aim*.) Commonsense psychology is home to several different forms of intentionality: it recognizes that beliefs have truth conditions, that desires have satisfaction or fulfilment conditions, that concepts have referents, and that perceptual states are *of* objects, properties, and states of affairs. These are all forms of psychological directedness—ways in which mental states are directed on things other than themselves. The objects of these sorts of directedness are not always real. The state of affairs that is the truth condition for a belief may not be actual, the state of affairs that would satisfy a desire may never obtain, the entity to which a concept refers may not be real (the concept of Hamlet refers to a fictional character), and an object that one perceives may be hallucinatory. But it seems likely that the objects of most intentional states exist. After all, most of our beliefs and perceptions are about trivialities, such as the present state of the weather, that we could not easily be wrong about. Also, the desires on which we act are largely realistic, and we generally have considerable evidence for the existence of the referents of the concepts that we most frequently deploy.

In 1911 the German philosopher Franz Brentano promulgated his *intentionality thesis*, which claims that intentionality is one of the properties that distinguish mental states from physical states of all kinds, including states of our bodies and physical states of the external environment (Brentano 1973). To be fully specific, he claimed that a state counts as mental just in case it possesses intentionality. This is a daring thesis. It isn't at all clear, for example, that moods like depression and boredom are directed on objects. But it seems to hold in enough cases that it makes sense to put mental phenomena such as moods under a reflective microscope, with a view to determining whether they may not harbor an intentional dimension after all.

Contemporary philosophy has responded to traditional discussions of intentionality, such as Brentano's, by endorsing the idea that the intentional states of the mind are grounded in mental representations, understood as causally interacting particulars that have veridicality conditions or aboutness conditions. It is thought that this idea articulates what is common to the intentional states posited by folk psychology, and also that it takes us part way to an explanation of their

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0001

directedness. Of course, the ultimate success of the idea depends on what we go on to say about the nature of representation. But assuming that we can develop an adequate metaphysics of representation, the idea promises to add considerable depth to our understanding of mental states. This is true no less of the mental states recognized in cognitive science than of the states recognized in folk psychology, for many of the states recognized in the former have their original home in the latter. Beliefs, desires, intentions, memories, emotions, and perceptual states are examples. When these states travel from folk psychology to cognitive science they carry their intentionality with them. This intensifies the need for an explanation of intentionality. What is it for a state to be directed on something? Arguably a theory based on representations holds the answer.

Quite apart from any role they may play in metaphysical accounts of the bases of psychological states, mental representations have come to figure explicitly in virtually all parts of contemporary cognitive science, ranging from discussions of place cells in the rat hippocampus, through accounts of higher level perceptual processing, such as accounts of how shapes can be recognized on the basis of shading, to explanations of mathematical expertise. Evidence for this claim quickly comes to hand as one pages through psychology and neuroscience journals, and it is found in a more compressed form in textbooks summarizing contemporary research in those disciplines (e.g. Palmer 1999, Wolfe et al. 2006, Frisby and Stone 2010, Yantis 2014, McBride and Cutting 2019, Kandel et al. 2021).

What is the reason for this strong, virtually universal commitment among cognitive scientists to representational theories of mental states? The first part of the answer is that all such theories are versions of information-processing psychology, which is broadly esteemed because it has truly extraordinary explanatory power (see, e.g., Palmer 1999, 59–85). The second part has to do with the limitations of *pure* information-processing psychology—that is, information-processing psychology that posits information-bearing states but not representational states. There are powerful reasons for thinking that pure information processing psychology must be supplemented with the concept of representation if it is to achieve the explanatory goals of psychological theories. Putting these two parts of the answer together, we get the result that representationalist psychological theories are the reigning champions because they are what you get when you elevate pure information processing psychology to a position of greater adequacy by supplementing it with the notion of representation. This additional notion enables information processing psychology to accommodate some important facts about mental states.

I will discuss three such facts. One is that perceptual states enjoy a specificity or determinacy of intentional content that pure information processing theories of perception cannot explain. The second is that perception can be mistaken or inaccurate. Pure information processing psychology has no room for talk of inaccuracy. Finally, there is the fact that perceptual states have a normative or

quasi-normative status that cannot be captured by any purely descriptive account of perception, including information processing psychology.

To illustrate the first point, suppose you see an apple tree that is situated with respect to yourself in a certain definite way (e. g., at a certain angle of view). The tree is projecting images on your retinas, but you do not see the images. We can express this by saying that the situated tree is part of the intentional content of your perceptual state but that the retinal projections of the tree are not. Now let us ask, can we account for these facts about intentional content in purely informational terms? It is hard to see how. There are two main ways of explaining information—in terms of causation and in terms of probability. On causal accounts of information, if a state S carries information about one of its causes, C, it also carries information about all of the other events that occur in the causal chain leading from C to S, and about all of the events that are causal antecedents of C. On such accounts, therefore, the information content of your current perceptual state would include both the apple tree before you and the retinal projections of the tree. Further, probabilistic accounts of information also fail to respect the specificity or determinacy of perceptual content. Suppose we say that the occurrence of an event or state E1 carries information about another event or state E2 if the occurrence of E1 makes it likely that E2 is also occurring—that is, if the conditional probability **P**(E2 is occurring/E1 is occurring) is high. Now your current perceptual state makes it likely that an apple tree is before you. It is also true, however, that your current state makes it likely that there are images of the tree on your retinas, where the images have aspects that correspond to aspects of the way the tree is situated. Hence, this probabilistic explanation of information implies that the information carried by your perceptual state includes information about the images. We get a quite different result, however, if we explain the intentional contents of perceptual states in terms of representational contents. Consider a painting of the tree. Any such image represents the tree, as situated relative to the painter, but is silent on events that count as causes or effects of the tree. And this holds in general: a representation can signify a state or event E without having any bearing on other states or events in the same causal chain as E, and without having any bearing on other states or events that are linked probabilistically to E.

Second, actions guided by perception often fail to achieve their goals. They are often unsuccessful. We must often explain these failures of action in terms of perceptual lapses—the actions fail because they are guided by inaccurate perceptions. Reflection shows, however, that pure information processing psychology has trouble accommodating inaccuracy. The fact that a state bears certain information is an essentially positive fact about the state. The information may be irrelevant to the task at hand, but irrelevance is compatible with being correct. Again, the information contained in a state may be too impoverished to enable an agent to complete a task successfully, but this is a limitation, not an error. To be sure, we

sometimes speak of *mis*information, but that usage is pretty much confined to cases in which speakers make false claims. It presupposes the idea of linguistic meaning, and therefore a concept of representation. In view of these considerations, we must seek beyond the informational properties of perceptual states to account for the fact that they can be erroneous. And representational properties seem to be just what we want. There is such a thing as representational error. Indeed we speak of *mis*representation all the time. Accordingly, if psychology assumes that the information-bearing states to which it is committed have the status of representations, it can extend its scope significantly, accommodating errors and thereby acquiring the ability to explain the many mental states and actions that are caused by errors.

Third, in addition to the dimension of perception that is captured by speaking of accuracy and inaccuracy, correctness and incorrectness, there is also a closely related normative or teleological dimension that is captured by talk of success and failure. When perception is accurate it is successful; it is doing what perception is supposed to. On the other hand, when perception is inaccurate, it is failing to do its job, thereby putting the organism at risk. There is a departure from how things are supposed to be, a failure to conform to a norm or standard of some kind. Moreover, the norm can't be purely causal or statistical in character. An event that departs from a causal or statistical pattern may be an anomaly but it isn't thereby a failure. (I am less than the average weight of American males, but this isn't a failure on my part.) As we have seen, however, our conception of information seems to be essentially causal and statistical in character. Talk of information carries no implications concerning conformity to norms or standards. On the other hand, talk of representational correctness or inaccuracy provides support for evaluations. Correct representation is a form of success, and misrepresentation is a type of failure.

I have suggested that these considerations are reasons to think that pure information processing psychology should be supplemented by a commitment to representations. One might suppose that they are instead reasons for supplanting information processing psychology with a new type of explanatory framework, but that would be a mistake. Representational states can carry information about the present environment, and in the case of perceptual representations, they normally do. Indeed, they could not otherwise perform their usual services for the organism. When information processing psychology is raised to a new level by adding the hypothesis that mental states are representational, it retains its basic commitments, including both its commitments to processes and its commitments to states. The change is just that the states are now regarded as having additional properties—properties that cannot be fully explained by the states' informational affiliations.

This brings us to the contemporary version of Brentano's intentionality thesis—the *representational theory of mind* (aka *representationalism*). This is the claim

that all mental states are constituted by representations and all mental processes are constituted by operations on representations.

Although the representational theory of mind currently holds sway in empirical science, and promises to unify and explain the intentionality of many of the mental states recognized in both science and folk psychology, reflection shows we are not yet in a position to accept it as true.

One reason is that we are not yet in possession of a sufficiently detailed version of representationalism to determine whether the theory can fully meet the explanatory challenges facing it. Can it give an adequate account of the phenomenology of perceptual experience? This question is very much open, as are many other questions about its explanatory adequacy. Another reason is that we do not yet have a satisfactory taxonomy of mental representations. This is largely because we do not have a satisfactory account of their formal or intrinsic properties. No doubt perceptual representations differ from conceptual representations, but exactly how? And how do representations of these two kinds differ from cognitive maps? To be adequate, a version of representationalism would have to answer such questions. A third reason for not committing fully to representationalism has to do with its universality. As I have defined it, representationalism claims that all mental states are representational. Reflection shows, however, that there are mental states that do not obviously have the sort of directedness that representationalism is needed to explain. The principal exceptions are pains and other bodily sensations, moods, and certain emotions, such as generalized anxiety (the form of anxiety that lacks a specific focus). A fourth reason is that the value of explanations in terms of representations depends on our being able to find plausible theories of the relation *X represents Y*. What is it for a mental particular to stand for an object, a property, or a state of affairs? Unless a satisfactory answer to this question can be found, explanations in terms of representations will come to seem quite shallow. I will discuss this last issue at length in the present chapter. The other issues are addressed later on in the book.

Before concluding this initial discussion of representationalism, I should remind the reader of an important distinction. Many representations are the result of natural causes. This is true, for example, of the representations that figure in perceptual processing. They possess representational contents that are not imposed on them by human plans, intentions, or conventions. But there are other representations that are brought into existence by human design and owe their contents to the human purposes and other mental states that lie in their causal background. Examples include paintings, maps, musical scores, and fuel gauges. It is customary to say that representations of this kind have *derived content* or *derived intentionality* because their contents depend on human plans or conventions for producing and interpreting them. It is important to keep this distinction in mind because doing so helps to fix ideas, but I will be focusing almost

exclusively, in the present chapter and in most of the others, on natural representations and natural contents.

II. Goals

The principal aim of this book is to expand our understanding of perceptual experience, but this overarching goal has five subordinate goals as components.

One is to advance the cause of representationalism by proposing detailed representationalist accounts of the main dimensions of perceptual experience. Such accounts should address questions about the format of experiential representations and also questions about their contents, where the contents of representations are the objects, properties, and states of affairs that representations represent. My proposals will offer answers to questions of both types.

In addition to the representations that constitute perceptual experience, there are also perceptual representations of other kinds. Some precede experience and others are post-experiential. These other representations will not be in the foreground, but at several points I will describe relations linking them to experiential states. I will also be concerned to explore questions about non-perceptual forms of experience. These other forms include bodily sensations, imaginings, and the conscious, qualitative aspects of moods and emotions. The proposals of a theory of perceptual experience should at least be consistent with plausible proposals about the nature of these other realms. I will sketch accounts of the other realms that are in keeping with my account of experience in the perceptual domain.

A second subordinate goal is to improve our understanding of the ancient problem of appearance and reality. Traditionally, discussions of this problem have tended to oscillate between two extremes. Many philosophers and psychologists have thought that if we are to do justice to the transitory and highly perspectival character of perceptual experience, we must suppose that the objects of direct perceptual awareness are entirely or largely subjective. The most familiar version of this view is the *sense datum theory*, according to which the direct objects of experiential awareness are mind-dependent entities called sense data. At the other end of the spectrum is *naïve* or *direct realism*. Shared by many philosophers and psychologists at the turn of the century, this view claims that the immediate objects of perceptual experience are mind-independent entities like shoes, books, trees, and armadillos, and that perceptual experience characterizes such objects by attributing objective physical properties to them—that is, properties such as objective sizes, shapes, and velocities. Advocates of direct realism tend to downplay, or even totally discount, the extent to which perceptual experience is transitory and perspectival. As I see it, neither of these theories is acceptable, and the same is true of most of the theories that lie between them. I will develop an alternative view, which draws a sharp distinction between experiential awareness of

objects and experiential awareness of *properties*. According to the view I will recommend, experience gives us direct access to external physical objects, but not to the objective properties of such objects. That is, we are not experientially aware of such properties as objective physical sizes, objective physical shapes, and objective physical distances. Instead we are aware of relational, viewpoint-dependent properties—that is, properties like visual angles. Appearances are constituted by properties of this kind. Now anyone who holds such a view must acknowledge that we also have some grasp of non-relational, objective properties of physical objects, for it is clear that the success or failure of our actions depends on our ability to negotiate environments that are defined by the objective natures of things. Accordingly, in addition to giving an account of perceptual experience that does justice to appearances, I will also offer an explanation of how perceptual experience combines with other factors, and in particular with cognitive faculties, motor programs, and post-experiential perceptual processing, to provide us with access to objective properties. Chapters 2 and 3 explain and promote these views about awareness of properties. Chapter 4 proposes a theory of how perception provides access to external physical objects.

The third of my subsidiary goals is to provide accounts of perceptual phenomenology and perceptual consciousness that are unified with what we know of the physical world. The phenomenological and conscious aspects of perceptual experience have always been among the most important platforms for arguments for mind/body dualism, the view that mental states are metaphysically independent of the physical facts of the world. It has been argued, for example, that the phenomenological dimensions of perceptual experiences are subjective, in the sense that they are only available from the first person perspective, whereas all of the physical bases of perception are objective, where this means that they can be fully grasped, in principle, at least, from the third person perspective (Nagel 1979, Jackson 1986). A closely related view is that consciousness is sui generis. There may be causal interactions between conscious experiences and physical phenomena in the world as we know it in everyday life, or in the world as it is known from the perspective of science, but the relationship between consciousness and physical phenomena is fundamentally contingent. There are domains of consciousness elsewhere in the actual world, or at least in other possible worlds, that have no physical tethers. Chapters 5–7 challenge dualistic claims of these kinds, and propose positive accounts of phenomenology and consciousness that fit comfortably within a physicalist perspective. One of the key ideas is the representationalist thesis that all awareness, including awareness of phenomenology, constitutively involves representations. This thesis makes it possible to maintain that awareness of phenomenology is governed by a Kantian appearance/reality distinction—a distinction between phenomenology as it is represented as being and phenomenology as it exists in itself. Careful examination of the intuitions and arguments supporting dualism reveals that they all presuppose the impossibility of drawing

an appearance/reality distinction with respect to phenomenology, and that they are therefore undermined by this representationalist thesis.

My fourth subordinate goal is to clarify the relationship between perceptual experiences and higher level cognitive states—-particularly, fact-positing states like belief, judgment, and knowledge. A number of investigators have maintained that perceptual experiences resemble cognitive states in that they are composed entirely or partly of concepts (McDowell 1996, Quilty-Dunn 2016). On this view, perceptual experiences are cognitive units in their own right. At the other extreme are the empiricist philosophers and psychologists who have argued that perceptual experiences are the principal components of cognitive states—that cognitive states are literally built out of perceptual material (Barsalou 1999, Prinz 2002). A third, more moderate view is that cognition and perception have proprietary representational codes, but that they can influence one another causally (Block 2022). I go along with this third view to the extent of thinking that cognition and perception *tend* to deploy different representational schemes, but I believe that cognition also frequently avails itself of a scheme that is closely related to the one involved in perceptual experience. Much of cognition makes essential use of perceptual imagery and perceptual simulations. Indeed, it is not implausible that animal cognition consists entirely of perceptual and quasi-perceptual representations. Still, this leaves us with the task of describing how exactly cognition differs from perception when they are using different representational schemes. This task occupies Chapter 8. Among other things, I maintain there that perceptual representation differs from conceptual representation in that it exploits certain types of isomorphism between representations and entities in the domains being represented. Then, in Chapter 9, I paint an externalist picture of how perceptual experience contributes to epistemic justification and knowledge. (Roughly speaking, epistemological externalism is the idea that whether a belief or a judgment is epistemically justified depends in part on circumstances that are external to the mind, such as whether the cognitive process that produced the belief has a tendency to produce beliefs that are true.) I will also maintain, however, that in addition to the externalist doxastic norms governing epistemic justification and knowledge, there are also internalist doxastic norms governing the doxastic virtue of *rationality*. At the end of Chapter 9 I will discuss some of the ways in which perceptual experience contributes to this virtue.

In addition to promoting externalist views in perceptual epistemology, and in epistemology generally, I will be trying to strengthen the case for externalist views in the philosophy of mind. A number of previous attempts to establish a general psychological externalism have run aground on perceptual phenomenology, which is widely regarded as internal. My fifth goal is to show that there is a tenable externalist account of this domain—that is, an account which maintains that the qualitative properties that constitute phenomenology exist independently of our awareness of them and independently of mental states and processes generally.

Pursuit of this goal will require a solution to what Adam Pautz calls the *external-internal puzzle about sensory properties*. Here is how he formulates the puzzle:

> Even in totally normal perception, and not just in illusion and hallucination, your experience of sensible properties (pain qualities, smell qualities, color qualities) is shaped by internal processing. Yet you experience sensible properties as "out there", together with shapes and in various locations, often at a distance from you. How is it that what you seem to experience as "out there" is shaped by internal processing "in here"? (Pautz, 2021, 9)

The solution I will propose is inspired by computational models of perceptual processing. It is a fundamental presupposition of such models that perceptual processes compute functions of perceptual representations—or, more accurately, functions of the contents of such representations. These computational processes are internal. Moreover, they determine what properties of objects we perceive. Hence, the properties we are aware of in perception depend on what happens "in here." I will maintain, however, that the functions computed by perceptual processes traffic only in certain properties of external objects that count as being "out there." More specifically, I will argue that the computed functions are concerned with properties of external objects that are relational and *viewpoint*-dependent but not *viewer*-dependent. The functions take properties of this sort as inputs and yield different properties of the sort in question as outputs. The trick consists in finding functions that answer to this general description and also have certain other desirable characteristics. Fortunately, as we will see, it is plausible that the trick can be performed. This line of thought is developed in Chapters 2 and 3.

III. Fixing Ideas about Perception and Experience

Since the nature of perceptual experience is the unifying theme of the book, I should say something about the nature of perception in this introductory chapter, and also something about the nature of experience. I will not attempt to give rigorous definitions. Rather, I will just make explicit some general principles that I believe to govern perception and experience, and will be presupposed in subsequent chapters.

First, with respect to perception, I will be presupposing that if a system of representations and processes is perceptual, then its representations stand for mind-independent entities. Second, I will assume that when a perceptual system is functioning properly, its representations and processes are to a large extent controlled by information about such entities. And third, I will assume that perceptual processing begins with the activity of sense organs—that is, with organs

whose function it is to pick up information about external objects and transduce it, converting it into signals that can be consumed by downstream perceptual processors. These assumptions are shared by a broad community of investigators—a community that includes both many philosophers and many psychologists.

It is more difficult to get a fix on experience, but there is one feature of experience that seems to be recognized by virtually all people who talk or think about the topic. This feature is consciousness. The principle that experiences are conscious seems to be baked into our conceptual scheme and therefore into the English language. Here, for example, are two entries found respectively in the *OED* and the *Merriam-Webster Collegiate Dictionary:*

> Experience (4): A state, condition or event that consciously affects one; the fact or process of being so affected.
>
> Experience (3): The conscious events that make up an individual life.

I will be assuming that experiences are conscious throughout the book. I will also assume that all *perceptual* experiences have a phenomenological dimension. I believe that this second principle is no less widely accepted than the first. Perceptual phenomenology consists of qualities like tasting sweet, smelling like cinnamon, feeling smooth, feeling warm, sounding high pitched, sounding loud, looking brown, looking spherical, looking small, and looking distant. It is, I think, widely acknowledged that all perceptual experiences involve properties like these—that is, ways of appearing to conscious observers (Kim 2006, 225). But what about conscious experiences that are not perceptual? It is less clear that they all have proprietary phenomenologies. Consider conscious thoughts. They are paradigm cases of experiences, but some philosophers have expressed doubts that they have a phenomenological dimension. Further, if conscious thoughts do have phenomenologies, how do their phenomenologies differ from perceptual phenomenologies? I will explore these larger issues about the realm of experience in Chapter 5. What matters for our present concerns is just that we get a fix on the nature of perceptual experience, for that is the main topic of this book. The thesis that perceptual experience has a phenomenological dimension seems adequate for this purpose, especially when that dimension is nailed down to qualities like the ones I have just listed.

But something more should be said about what those qualities have in common. They can all be described as ways in which objects appear to us perceptually (Kim 2006). But what is the meaning of the umbrella term "appear" in this context? To answer this question, I need to invoke Roderick Chisholm's observation that "appear" and its cognates have three different senses (Chisholm 1957). In what Chisholm called the *epistemic sense* of "appear," an object O may be said to appear to a subject S to have a property F just in case S believes that O has F and

this belief is supported by evidence that is possessed by S. In the case of a perceptual belief, S's evidence would consist of one or more perceptual experiences. The second sense Chisholm distinguished is the *comparative sense*. When "appears" is being used in this sense, an object O may be said to appear to a subject S to have F just in case O appears to S the way things that have F normally appear when they are perceived in standard or optimal circumstances. Chisholm's example of this second sense was the claim that railway tracks appear to converge as they recede into the distance. Although there are other ways of understanding this claim, it can be taken to mean that tracks appear the way that genuinely convergent lines appear when they are orthogonal to the line of sight. This brings us to the *non-comparative sense* of "appear," which is the one that is presupposed above and will be presupposed elsewhere in the book. This is the sense we have in mind when we just mean to be describing an experience, without mentioning its epistemic role or its similarity relations to other experiences. Is there such a sense? It is clear that it is possible to speak of appearances without claiming that they provide epistemic support of any kind, whether for beliefs or simply for propositions that a subject can entertain. Thus, for example, one might say correctly that an object appears hexagonal to a chimpanzee, even though the chimp lacks the concept of a hexagon and is generally unable to entertain propositions about complicated shapes. But what about the comparative sense of "appear"? Can we use it to explain the non-comparative sense? In other words, when we say that O appears to S to have the property F, could this just be shorthand for the claim that O appears to S the way things that have F normally appear when they are perceived under optimal conditions? There are strong reasons for doubting any such reductive claim. For instance, as James Van Cleve has pointed out (personal communication), the reducibility thesis leads to an infinite regress: "In the analysis 'x looks red to S =df x looks the way red things normally look to S', there is a reference to a way of looking on the right. If that way must itself be understood comparatively, we are off on a regress."

Chisholm's threefold distinction nicely captures differences that I have noticed in my own thought and talk about appearances, and that I have also noticed in the discourse of others. Accordingly, I will assume that there is a non-comparative sense of "appear" that can be used to express a general thesis about perceptual phenomenology—specifically, the thesis that the phenomenological dimension of perceptual experience consists of perceptual appearances. This thesis will be one of the principal topics of Chapter 2.

IV. Teleosemantics Plus Representational Pluralism

The achievements and promises of information processing psychology are so great that it is hard to imagine life without it. As we saw, however, there are

powerful reasons for thinking that *pure* information processing psychology is inadequate: to achieve completeness, it must be supplemented with the concept of representation, and this concept must be applied generously, so that it figures in the characterization of all perceptual states.

But what are we adding when we suppose that perceptual states have the status of representations? There have been a number of attempts to answer this question, including efforts by Stampe (1977), Dretske (1981, 1986, 1995), Millikan (1984, 1989a) Dennett (1987), Papineau (1987), Cummins (1989, 1996), Fodor (1990, 2008), Bermudez (2003), Burge (2010), Neander (2017), Shea (2018), and Williams (2019). I will focus here on the family of views that count as versions of *teleosemantics*. Roughly speaking, teleosemantics seeks to analyze perceptual representation in terms of the concept of information and the concept of a function. It asserts that perceptual states have informational functions, and that it is this that confers the status of representations upon them.

Beyond this very broad characterization, it isn't easy to identify a set of commitments, or even a single commitment, that all teleosemanticists share. What can be said, though, is that versions of the following principle are at least widely accepted by members of this group:

(TS) A perceptual state S represents a property P only if it is a function of S to encode information about P.

If we remember that there are a number of different ways of defining information, and also a range of alternative accounts of functions, we can appreciate that (TS) captures the common core of a number of independent views. But still, it is sufficiently concrete to illustrate some of the main benefits of teleosemantics.

As we saw, representations can be erroneous. (TS) nicely accommodates this fundamental feature of representations. It can be the function of a state to encode information about an entity even though it is not currently performing that function. This is just a special case of a general property of functions. Thus, it is the function of kidneys to extract wastes and toxins from the blood, and it is the function of a fuel gauge to carry information about the level of available fuel; but due to age or damage, these mechanisms may fail to perform their respective services. In general, there is a gap between having the function to perform F and actually performing F. Because of this gap, if perceptual states have informational functions in the sense captured by (TS), we can understand how they can be erroneous by saying that they are failing to perform their informational functions.

To elaborate, when an event is described as an error, there is a suggestion that there has been a failure of some kind—the event is somehow flawed. But the notion of a function is tightly connected with the concepts of failure and flaw. It is common to describe kidneys and fuel gauges as failing to perform their functions, and to explain these failures in terms of disorders, breakdowns, and structural flaws.

This quick line of thought illustrates the appeal of theories of representation that imply (TS). But some authors would say that (TS) is much too strong to be true. As an example of this sort of view, consider a complaint about principles like (TS) that was once lodged by Ruth Millikan (Millikan 1989a). It has to do with the notion of information. On one account of the nature of information, a state or occurrence X cannot carry information about a condition C unless the conditional probability **P**(C/X) is extremely high (Dretske 1981). But, as Millikan pointed out, if we construe information in terms of high conditional probability, (TS) will conflict with obvious facts about representation: many things in nature count as representations even though they do not confer high probabilities on the conditions they represent. For example, the slap of a beaver's tail on water is a representation of danger. But beavers frequently slap their tails when no danger is at hand, with the result that the conditional probability **P**(danger/tail slap) is far from high. On the face of it, it follows that beaver behavior is a counterexample to (TS).

In fact, however, (TS) is sufficiently resilient to meet this challenge. It is true that there is a conception of information on which information entails high conditional probabilities. But there is another conception on which it just involves *probability raising*. Thus, on this alternative conception, a state or occurrence X carries information about a condition C just in case **P**(C/X) is significantly higher than **P**(X). On this weaker conception of information, (TS) can easily accommodate the representational properties of tail slaps.

This illustrates the flexibility of (TS): more than one account of representation can fit under its umbrella. Indeed, a number of accounts can fit, for in addition to the two conceptions of information we have just canvassed, there are other accounts that explain information in terms of relations of covariation between two variables, rather than relations of conditional probability (Cummins 1989), and still other accounts that seek to explain information in terms of causation (Fodor 1990, Fodor 2008, Neander 2017). On the latter views, there is a type of information I such that a state or occurrence X can only carry I-information about a condition C if C is a reliable cause of X.

We need not see these conceptions of information as competing with one another. Instead of thinking of them as contending accounts of a single phenomenon, we can view them as singling out different relations that are united by family resemblance. By the same token, it seems legitimate to adopt a pluralist view of teleosemantic theories of representation. On this pluralistic view, we can get equally viable accounts of representation by plugging different conceptions of information into (TS). To be sure, it might be that one or more of these accounts has little to offer cognitive science, but that is not a priori true. It may well be that different branches of science will find different accounts useful, and/or that different accounts will prove suitable for different levels of explanation. Indeed, on reflection this seems likely. It is plausible that theorists would want to distinguish

between different grades of representational reliability and different grades of representational stability. One natural way of achieving this is to recognize different types of representation, where the types are individuated by the different probabilistic and causal relations by which representations are linked to their respective representata.

The merits of pluralist accounts of representation come to the fore when one reflects on the question of whether perceptual states represent properties that correspond to biological needs, such as the property of being food, or only represent more basic properties, such as sizes, shapes, and forms of motion, that are *correlated* with need-answering properties in an organism's biological niche. It will be helpful to have a concrete case of this general question before us, so let us consider what the representational content of a certain type of activity in the visual system of frogs—all it *activity of type A*—might represent. A-activity occurs whenever a small black object moves across a frog's visual field. Should this lead us to suppose that it represents the property *small, moving, black object*? Perhaps, but it is also true that the small, moving, black objects in a frog's biological niche have a strong tendency to be living creatures, and therefore a strong tendency to be food. Hence, when it comes to fixing upon the representational content of A-activity, we have at least two choices, one being that it represents *small, moving, black object*, and the other being that it represents something like *food* or *nutrient.* There is something to be said for both of these possibilities. The first hypothesis is favored by the fact that A-activity is more strongly correlated with the first property than the second. Guided by A-activity, frogs will extend their tongues to capture small, moving, black dots even if the dots are bee-bees that have been painted black and are being moved by threads. This might suggest that the first property is the one that is the actual representational content of A-activity, but there is also another side of the story. After all, it is the property *food* that answers most directly to a biological need. Shouldn't properties with ecological value be the ones that we take to be contents? After all, it is plausible that Mother Nature has taken the trouble to endow frogs with visual systems only because they can effectively guide actions aimed at goals that are set by needs.

We are confronted with many similar choices. For instance, as Karen Neander has pointed out (Neander 2017), there are cells in the visual systems of toads (let's call them *B-cells*) that fire whenever they are stimulated by objects that are small, rectangular, and moving along a line that is parallel to their longest axis. That is, B-cells fire whenever a stimulus has these three properties, even if it is a piece of cardboard being dragged along the ground by a lab assistant. Given the tight causal relations between stimuli with these properties and activity in a toad's B-cells, it can seem that there is a compelling case for the idea that the former figure in the representational content of the latter. But we are given pause when we learn that most of the small, rectangular, moving stimuli in a toad's niche are

insects or worms, and therefore have the status of food. This suggests that it is a property like *food* that figures in the content.

It is often controversial how dilemmas of this sort are to be resolved. Thus, the dilemma involving frogs sparked a fracas between Millikan and Fodor (Millikan 1986, Fodor 1990), and Neander reports that there is no settled resolution of the dilemma involving toads within the community of toad scientists. But perhaps debates about such issues rest on a false presupposition. It is assumed on both sides of such controversies that there is only one scientifically relevant type of representation, and that the goal of theorizing about representation is to characterize that one type. When we reflect, however, we realize that it may be possible to take a pluralistic attitude toward questions of representation in science. Indeed, reflection suggests that pluralism is the *right* attitude. Since there is a strong case for supposing that perceptual representations stand for the low level properties that causally trigger them, and also a strong case for supposing that they stand for the properties that are topics of ultimate biological concern, there is a strong case for supposing that perceptual states can stand in representation relations of two different kinds. This case would be defeated if there was a compelling argument for supposing that perceptual representation is unitary, but, as far as I know, no argument of this sort has ever been given and it is hard to imagine what form such an argument might take. Moreover, it is easy to see how scientists of different stripes might prefer to focus on different forms of representation. Scientists who are interested in the causal mechanics of the visual system would presumably be primarily interested in a form of representation that linked perceptual states to low level properties like *small* and *rectangular*, while scientists with an ecological perspective would be interested in a form that linked perceptual states to properties like *food* and *mate*.

In speaking of different scientific perspectives, I do not mean to suggest that the different forms of representation are in any sense dependent on or relative to these perspectives. The view is rather that there are multiple relations of representation, all existing independently of human explanatory perspectives and modes of inquiry, but that different members of this class of relations are of interest to different groups of scientists. The point of the preceding paragraph is that one of these groups focuses on the relation between representations and properties of distal objects that causally explain how the representations are triggered, while another group focuses on the relation between representations and properties of distal objects that explains how the representations serve biological needs. I add that these groups have subgroups with more specific agendas.

There is a question here in that it seems much more likely that perceptual mechanisms are sensitive to properties like *small* than to properties like *nutrient*. This becomes apparent when we recall two facts that we noticed a moment ago. A-activity occurs *whenever* a frog is confronted with a small, moving, black

object, and B-cells fire *whenever* a toad spots an object that is small, rectangular, and moving along a line that is parallel to its longest axis. It doesn't matter whether stimuli of these kinds are nutritious or not. This suggests that the property of being nutritious has no causal powers with respect to the perceptual states of frogs and toads. Is this compatible with thinking of those states as representations of the property?

We already have an answer to this question. As we observed a few paragraphs back, when it is combined with representational pluralism, teleosemantics can accommodate conceptions of information that are based on probability as well as conceptions that are based on causation. Now the conditional probabilities **P**(*nutrient/frog state A*) and **P**(*nutrient/toad state B*) are no doubt reasonably high, at least in the relevant biological niches, though the corresponding conditional probabilities for properties like *small* and *moving* are higher. We ought to conclude, then, that teleosemantics sustains the idea that ecologically significant properties can be represented perceptually. Another endearing consequence of representational pluralism!

These reflections can be generalized. Suppose that Σ is a system of perceptual representation, and that the individual representations that compose Σ have the function of encoding information about the properties in a domain D. Suppose also that the relevant form of information is causal—that is, the mechanisms that control the representations in Σ are causally sensitive to the members of D and to no other properties. We need two more assumptions. First, suppose that there is another domain D* such that the members of D indicate the members of D* with a reasonably high probability. It follows of course that the representations in Σ are also linked probabilistically to the members of D*. Finally, suppose that Σ was selected by evolutionary pressures because its constituent representations stood in these mediated probabilistic relations to the members of D*. (To illustrate, in an especially simple case, the sole member of D is the complex property *small, black, and moving* and the sole member of D* is the property *food*.) Given all of these assumptions, I suggest, it is appropriate to say both that the members of Σ represent the members of D and that they represent the members of D*. These representation relations are different in two ways. (i) The first relation is *mechanically* more fundamental than the second in the sense that the components of Σ could not encode probabilistic information about D* if they did not carry causal information about D, though they could encode causal information about D even if they didn't encode probabilistic information about D*. (Witness the cases of A-activity in frogs when they are shown painted bee-bees that are suspended from threads.) And (ii) the second relation is *ecologically* more fundamental in the sense that Σ would not exist if it was not probabilistically attuned to D*. (If its components hadn't been correlated with nutrients, Σ would never have evolved.) Despite these differences, however, the two relations count intuitively as forms of

representation, and they are both countenanced as such by selectionist teleosemantics. Thus, it can be true both that the members of Σ were selected because they encoded (causal) information about the members of D and that they were selected because they encoded (probabilistic) information about the members of D*. These two selectionist theses are not in competition with each other. Accordingly, it can be true both that the members of Σ have the function of encoding information about the members of D and that they have the function of encoding information about the members of D*.

There have been a number of controversies as to whether perceptual contents are *low* or *high* (*thin* or *rich*) that are more complicated than the ones about the perceptual representations of frogs and toads. The present dualistic account of representation can help to resolve them. Perhaps the most interesting of these controversies has to do with the system that enables human beings to discriminate and recognize the faces of other humans. Do we recognize faces as such or only congeries of lower level properties and relations such as shapes of parts, distances between parts, and colors of parts? A number of philosophers and psychologists have taken sides on this question, but, as I see it, both groups have been laboring under a false presupposition. They have assumed that the contents of facial representations must be either high or low. It is plausible, however, that pluralism is the right attitude to take toward this issue. Thus, it is plausible that perceptual states of a certain type Σ were selected because they encode probabilistic information about faces. Natural selection would not have bothered to endow us with Σ if this had not been true. Our survival depends on our having the ability to recognize human faces as such and the ability to discriminate among different faces. It is also plausible that the members of Σ were selected because they encode causal information about certain congeries of low level properties that are strongly correlated with faces (i.e. the shapes and sizes of components of faces, and spatial relations among those components). This is because the states in Σ could not have encoded probabilistic information about faces if they had not encoded causal information about those congeries of lower level features. Moreover, these two selectionist stories are compatible. They make different selectionist claims, but there are countless examples of cases in which pairs of selectionist stories are both true. For example, it is true that the adrenal glands were selected because they secrete adrenalin, and it is also true that they were selected because they help to prepare the organism for a fight-or-flight response to danger. In view of all this, it is natural to conclude that the members of Σ have two informational functions—the function of encoding causal information about D and the function of encoding probabilistic information about D*. Further, when it is combined with (TS), this conclusion supports the additional conclusion that the members of Σ represent members of D and also members of D*. (See Green 2017a for a complementary discussion of causal and ecological motivations for assigning contents.)

A nice illustration of the latter conclusion is afforded by the work of Doris Tsao (Chang and Tsao 2017, Tsao 2019). Tsao has shown that every human face corresponds to a point in a fifty-dimensional quality space, where the dimensions of the space are determined by low level magnitudes such as degrees of curvature of lines. Importantly, points in the space are registered by firing patterns in small neural nets. Now it is natural to suppose that these firing patterns were selected because they encode information about points in the space. But it is equally natural to suppose that they were selected because, in responding to information about a point in the quality space, they were thereby indicating the presence of a human face. (A presented item corresponding to a point in the space could be a lifelike mask, but the probability that it is actually a face is extremely high.) Moreover, both of these suppositions can be true. There is no need to choose between them. By the same token, there is no competition between the hypothesis that the firing patterns represent points in the space and the hypothesis that they represent faces. (I will revisit Tsao's brilliant work in later chapters.)

I should add that pluralism does not entail promiscuity. Thus, for example, although flying insects are frog food, and also possess the properties that serve as causal triggers for representations in Σ, the property *flying insect* does not imply either the ecological property *food* or the causally efficacious property *small, black, and moving*. Accordingly, even from the perspective of representational pluralism, there is no basis for saying that it is represented.

We have been considering ways in which the theory consisting of (TS) and semantic pluralism can be used to settle controversies about perceptual representation. Clearly, the theory has a lot going for it. But, of course, it also leaves many questions about representation open. Perhaps the most important of these is the question of what a sufficient condition of representation should look like. (TS) spells out an important necessary condition of representation, but it is silent on the question of sufficiency. I will not attempt to resolve this issue here. Instead, I will just emphasize that (TS) offers a foundation on which multiple theories of sufficient conditions can build, and that representational pluralism encourages us to expect that more than one of these theories will turn out to be correct.

These points are illustrated by a proposal about sufficiency that is suggested by Tyler Burge's path-breaking discussion of perceptual representation (Burge 2010). As I read him, Burge would accept (TS), or would at least accept a thesis in that neighborhood. But he would say that (TS) is far from capturing a sufficient condition of representation. Thus, according to Burge, it is a constitutive fact about perceptual representation that it puts us in touch with the objective properties of external objects. It is easy to see, however, that the necessary condition cited by (TS) can be satisfied by states that fail to do this. That condition is satisfied by any brain states that have the function of registering the transitory impacts of chemical, mechanical, and other stimuli on the surface of the body. Accordingly, Burge maintains, we must recognize that perceptual representation only occurs at higher levels of visual processing, after constancy transformations have produced

perceptual states that reliably covary with objective properties. That is, as Burge sees it, constancy transformations are necessary for perceptual representation. (Constancy transformations are discussed at length in Chapter 2. Roughly, they are information processing procedures that prescind from the constantly changing and highly perspectival properties of peripheral stimulation by performing computations on sensory inputs of two or more different kinds (for example, by integrating retinal inputs pertaining to size with retinal inputs pertaining to distance).)

Clearly, there is a lot to be said for Burge's claims. Suppose for the moment that we decide to accept these claims while continuing to accept (TS). That would give us a complex principle that looks like this:

(TS+) A perceptual state S represents a property P only if (i) it is a function of S to encode information about P, and (ii) the processes that are causally responsible for S include constancy transformations.

Now if we accept (TS+), we will want both clause (i) and clause (ii) to figure in any sufficient condition for perceptual representation we might propose. The intuitive appeal of Burge's claims provides motivation for theories of representation of this sort.

Interestingly, however, there is also motivation for a sufficient condition that does not involve clause (ii). Thus, while Burge thinks that perceptual representation requires constancy mechanisms, and has good reasons for his opinion, the dominant view in perception science appears to be that taste and smell are forms of perception, even though they show little influence by constancy transformations. Perception scientists are more impressed by the very substantial similarities between these two modalities and paradigmatic forms of perception like vision and hearing than by the differences. (Among other things, they have the function of encoding information about entities that are external to the mind, albeit entities that are in contact with bodily tissues. Additional similarities are discussed in Section VIII of Chapter 3.) Differences of opinion of this sort can be quite confusing. It appears, however, that liberal doses of representational pluralism can go a long way toward resolving them. Burge has good reasons for his views and the same is true of the community of perception scientists. It is therefore natural to suppose that they are discussing different forms of representation. This suggests that necessary and sufficient conditions for perceptual representation can be derived from (TS) in at least two different ways—one that involves adding a reference to constancy transformations, and another that involves adding a reference to states whose job it is to encode information about extra-mental entities. Both of these ways of proceeding are well motivated, and representational pluralism allows that both can be correct.

To summarize: When it is combined with representational pluralism, (TS) captures an idea that is shared by a number of different versions of teleosemantics,

and it resolves conflicts about the nature of perceptual representation in appealing ways. It therefore challenges what I believe to be a common impression concerning the teleosemantics literature, which is that it is a locus of Boschian chaos, a darkling plain swept with confused alarms of struggle and flight. As against this view, it seems reasonable to see the literature as containing strong indications of convergence and progress, while also offering a richly diverse set of ideas about different though related topics. In addition, it seems to provide a strong foundation for theories of perceptual representation that propose conditions for perceptual representation that are both necessary and sufficient.

I will henceforth understand teleosemantics to be the set of principles that contains (TS) and its various well motivated interpretations, together with all well motivated proposals concerning necessary and sufficient conditions for different types of perceptual representation that are based on (TS). When it is so understood, there is a case for using teleosemantics as a foundation for theories of perceptual representation that is prima facie quite strong. In any case, this is how teleosemantics appears to me, and I will therefore presuppose it in later chapters. To be specific, I will presuppose that it is the best way of implementing a representationalist approach to perception that is currently available. This is not to say, however, that the representationalist approach stands or falls with teleosemantics. The justification for representationalism that is given back in Section I is neutral among an extensive range of different metaphysical accounts of representation, and additional neutral justifications will be presented in later chapters. I believe the reader will find that most (though definitely not all) of the main claims of the present book would survive transposition into the vocabularies of those alternative accounts.

V. Selectionist Teleosemantics, Learning, and Cummins Teleosemantics

I turn now to the question of how we should understand (TS)'s use of the term "function." Like "information," this term can be understood in a number of different ways, and several of these ways have figured in the literature on teleosemantics. It is clear, however, that there is a widespread interest, on the part of teleosemanticists, if not a widespread commitment, to regarding functions of the relevant sort as belonging to the same category as biological functions—that is, functions of the sort we invoke when we say that it is a function of the kidneys to clean the blood, or that it is a function of the hypothalamus to trigger the production of hormones. Further, there is a widespread interest in giving a *selectionist* account of biological functions. In any such account, it is maintained that it is the biological function of entities of type T to do F if entities of type T were selected

by evolutionary processes because they did F. (Here entities of type T might be states, processes, or structures (organs, assemblies of neurons, etc.).) In more technical terms:

> It is a/the proper function of an item (*X*) of an organism (*O*) to do that which items of *X*'s type did to contribute to the inclusive fitness of *O*'s ancestors, and which caused the genotype, of which *X* is the phenotypic expression, to be selected by natural selection. (Neander 1991, 174)

To summarize, there is considerable interest in a two-part package that involves (i) interpreting the term "function" in principles like (TS) as referring to biological functions, and (ii) explaining biological functions by giving a selectionist account. As we will see, this proposal is not without problems, but it also has some impressive virtues. Thus, since it is possible for an item to possess a biological function F even though it is not currently performing F, the package supports (TS)'s ability to account for representational failure, as in cases of illusion and hallucination. Another virtue is that in combination with (TS), the package grounds perceptual representation in factors that do not themselves involve representation or intentionality. Not all accounts of function are free from this sort of circularity. Consider the following alternative account: It is the function of an item X of organism O to do F if doing F tends to promote the fulfilment of O's desires. Unlike etiological accounts of functions, which invoke developments in the history of a species, this *individualist* account implies that possession of functions depends only on dispositional facts about a single agent. Unfortunately, whatever attractions the account may have, it is of no use to anyone seeking a solid foundation for talk of representation and intentionality, for in appealing to desires, it invokes intentional states.

We can appreciate another strength of the package by comparing it to theories that seek to explain functions purely in terms of causal roles. On Robert Cummins's account of functions, the function of an item is its disposition to contribute causally to the normal operations of the larger system in which the item is housed (Cummins 1975, Ariew et al. 2002). There is no restriction on the types of causal contribution nor on the larger systems that contain the items. Accordingly, Cummins's theory faces a serious problem of overgeneralization: for example, as Millikan pointed out, it implies that it is "the function of clouds to make rain with which to fill the streams and rivers" (Millikan 1989b, 294). Perhaps Cummins can get over this problem by introducing appropriate restrictions, but there is no escaping from the objection that his theory cannot distinguish between items that are performing their proper functions and damaged or worn out items that are not performing their proper functions but are nonetheless making systematic contributions to the outputs of the system in which they are housed. The latter are

performing Cummins-functions, but they aren't performing functions of the sort recognized in biology. For related reasons, Cummins-functions cannot underwrite the concept of a perceptual error. In an unfavorable environment, it might be that most tokens of a perceptual representation are erroneous. But they could still be described as performing Cummins-functions, for they would be making contributions to the output of their housing system that were statistically normal. (Godfrey-Smith (1993, 200) maintains that there is room for a normative element in Cummins's theory in the following passage: "If a token of a component is not able to do what other tokens do, that plays a distinguished role in the explanation of the capacities of the broader system, then that token is malfunctional." But "not doing what other tokens do" just means that the given token is statistically aberrant. Deviations from statistical patterns do not count as errors unless it is also true that they are due to malfunctions.)

So far so good, but we must now consider an important objection to *purely* selectionist accounts of representation that has been emphasized by Fred Dretske and Nicholas Shea (Dretske 1988, Shea 2018). It is plausible that a large proportion of our perceptual representational capacities is due to natural selection. This is true, for example, of many aspects of our abilities to represent basic properties like sizes, shapes, distances, colors, and forms of motion. These abilities are part of our biological endowment. But it is also true that we have many representational abilities that are due to perceptual learning. To illustrate: we can learn to recognize a mountain from many different vantage points, and we can learn to discriminate between cedars and pine trees of other kinds. These abilities involve representations that are not products of natural selection. Hence, at first sight, anyway, if we wish to maintain that (TS) is a fully general principle about perceptual representation, we must look for a broader interpretation of the notion of an informational function than the one that is provided by the foregoing selectionist account. On that account, a state has an informational function only if it was selected by evolutionary forces as a result of its history of encoding information. Informational functions due to learning do not always fit this picture.

On the face of it, then, learning poses a serious problem for purely selectionist accounts of informational functions. Selectionist accounts offer plausible sufficient conditions for the existence of such functions, but it can seem that the conditions they propose are not necessary. I acknowledge the prima facie force of this problem. I believe, however, that it turns out in the end to be tractable. I will mention five considerations that address particular concerns about learning, and will then propose a more unified way of confronting the problem.

First, the scope of perceptual learning is quite limited: it cannot support learning that exceeds our perceptual capacities. Thus, for example, perceptual learning cannot give me the ability to see Rotarians as such. Nor can it enable me to distinguish Democrats from Republicans. This distinction is not associated with any

distinction between groups of perceptible properties. To be sure, it can be appropriate to say that someone sees that a Republican is present, or that a Democrat is present, but that can mean only that the person applies the concept of a Republican, or the concept of a Democrat, to someone who is seen.

Second, perceptual learning often involves no more than actualizing an inherited capacity to encode information—a capacity provided by natural selection. This is true, for example, in typical cases of *classical conditioning*. Pavlov noticed that a dog began to salivate when it heard the footsteps of a laboratory assistant coming to feed it. The sound of the footsteps had become a conditioned stimulus for salivating. But although the sound of the footsteps had this new role, the dog's ability to register the footsteps and encode information about them was not new, even if it had never been engaged in any earlier contexts.

Third, there is much discussion of learning by *Bayesian updating* in the literature on perception. This is a form of learning, but it doesn't call for representations that are genuinely new. It just involves redistributing probabilities over representations that are already available. (The reader will find an accessible account of Bayesian updating in Joyce 2003.) A similar observation applies to what is often called *reward learning*. Reward learning is principally concerned with updating value assignments to certain antecedently existing representations and strengthening or weakening connections between others. (There is a reader-friendly account of this form of learning in Arpaly and Schroeder 2014.)

Fourth, the operations that produce new representations often do no more than perform simple operations on representations that are already available, either potentially or actually. This is what is accomplished by two operations that figure prominently in the literature on perceptual learning—the operation of cobbling together previously existing representations (mechanisms of *unitization*), and the operation of subjecting unified representations to fission (mechanisms of *differentiation*). (For extended discussions of unitization and differentiation see Goldstone and Byrge 2015 and Connolly 2019.) Learning that is due to these mechanisms does not call for a significant expansion of accounts of perceptual representation that are based on (TS) and a selectionist account of information functions. Thus, for example, if each member of a set of states has an informational function, then it's not much of a stretch to suppose that the state produced by cobbling those states together will inherit an informational function from them, and will therefore come quite close to fitting the (TS)/selectionist profile. Indeed, this is the most natural view of the matter.

To be sure, once the pre-existing representations have been bound together, there is a new representation that stands for a property that is probably not explicitly represented by any representation already in the relevant agent's repertoire. But the potential for producing the representation was contained in the agent's prior representations and the agent's inherited apparatus for producing new representations by unitization. In this sense, the new representation was implicitly

contained in the agent's prior representational capacities. It seems perfectly natural to say that it has a derived representational function—a function that derives from actualizing a potentiality. (Similar remarks apply to new representations produced by differentiation.)

Fifth, it is possible to generalize this notion of a *derived* or *inherited* function. Once generalized, it can be used as a basis for explaining representations acquired by more powerful forms of perceptual learning, such as learning to pick out instances of a complex pattern by attending to its parts. Let M be a mechanism in a biological system. I suggest that it is appropriate to say that M has the derived function of performing activity A, where A is a kind of activity of the more general kind K, just in case M is the product of an inherited mechanism M* such that M* has the function of producing mechanisms that can perform activities of kind K. Applying this notion, it would be possible to say, for example, that a mechanism M has the derived function of registering information about a pattern P that is too complex to be obtained from simpler and more basic patterns by simple operations. For this to be true, it need only be the case that M is a product of a mechanism that has the function of producing mechanisms that register information about patterns in a family to which P belongs.

Is this proposal well motivated? Do we have any use for the notion of a derived function in other contexts? Yes. Indeed, the notion of a derived function is widely applicable. Consider an *inherited* mechanism M that performs A, where performing A counts as a function in the sense underwritten by the selectionist account of biological functions. Suppose M is a mechanism of the organism O. Now, on the selectionist picture, M has performing A as a function in virtue of the fact that M is a product of the genes that are responsible for O's inherited equipment. That is, M has the function of performing A because those genes had the function of producing a mechanism that could perform A. If M hadn't been produced by genes with that function (as would have been the case if M had been produced by a random process, or if the genes in question had been produced by a random process), it wouldn't have possessed the function to perform A. In a certain sense, then, *all* of the basic biological functions of the states and organs of an organism are derived, for the states and organs that possess these functions are produced by genes that have the function of producing items that behave the way those states and organs behave, and they would not have had those functions if they had not been produced by such genes.

It turns out, then, that there are a number of things that selectionists can say in response to the challenge posed by perceptual learning. Even so, the fact that there is this prima facie serious challenge suggests that the selectionist interpretation of (TS) may be too narrow. The domain of perceptual learning is vast and diverse. Perceptual learning can be automatized and it can also be under voluntary control. It can take the form of changes in the weights on connections among

neurons in early vision, and it can consist in adjustments in the evidential values of full blown perceptual experiences. It can involve improved discrimination and recognition of (i) single features, (ii) complex conjunctions of features, or (iii) higher order patterns that entail complex relations among objects or features (such as the relations among parts of faces). It can be fast or slow, and its effects can be enduring or ephemeral. Some forms of it require attention, while others do not. And so on. (For a comprehensive survey of these and many other dimensions of perceptual learning, see Dosher and Lu 2020.) To be sure of accommodating this large and variegated territory, it seems best to broaden the foundation of (TS) by complementing the selectionist interpretation with another one. That is what I will do now.

As we saw, Cummins-functions are properties of states or parts of systems. They owe their status as functions to the causal contributions they make to the processes that mediate between the inputs and outputs of the systems. They can be possessed by states and parts of anything that qualifies as a system, and they can be exhaustively explained in causal and statistical terms. Because of these features, they cannot explain quasi-normative features of representations such as accuracy, veridicality, and error. But let us now focus on the comparatively small subset of Cummins-functions that are possessed by states and parts of biological organisms, and more particularly, on the subset of Cummins-functions that make contributions to the proper functioning of the organisms in which they are housed. In these biological contexts, with their proprietary teleological dimensions, Cummins-functions have a new and more elevated status. Thus, consider a state S of an organism O that normally encodes information about a property P. Let it be true that S contributes to the proper functioning of O by playing this informational role. Suppose now that a token of S occurs even though there is no instance of P at hand. This token of S fails to contribute to the proper functioning of O, so we have a lapse that is more than a deviation from a statistical norm. It seems fair to say here that an error has occurred, for the token could cause serious problems for O.

There is a strong case, then, for the view that *teleological Cummins-functions* can ground the quasi-normative aspects of representation. Assuming this is right, it is appropriate to interpret the reference to functions in (TS) in terms of informational functions that are defined as follows:

A state S of a system Σ has the teleological Cummins-function of encoding information about a property P just in case (i) Σ is a biological system, (ii) S has the capacity to encode information about P, (iii) in virtue of having this capacity, S contributes to the output of Σ (or is poised to do so), and (iv) in making this informational contribution to the output, S is contributing to the proper functioning of Σ.

In short, an interpretation of (TS) in terms of teleological Cummins-functions is no less suitable than an interpretation in terms of selectionist functions. This is helpful because states with teleological Cummins-functions need not owe their existence directly to selective pressures. They can meet the defining conditions of such functions even if they result from ontogenetic factors.

These considerations suggest a bifurcation of (TS) into two principles that respectively deploy the selectionist conception of a function and the teleological Cummins conception:

(TSD) A perceptual state S represents a property P *directly via natural selection* (that is, S *D*-represents P) only if it is the selectionist function of S to encode information about P.

(TSI) A perceptual state S represents a property P *indirectly via natural selection* (that is, S *I-represents* P) only if it is the teleological Cummins-function of S to encode information about P.

What we are observing here is another advantage of combining (TS) with a pluralistic attitude toward perceptual representation. As the reader already knows, it is my view that a number of the main quarrels among teleosemanticists, and a number of the main objections to teleosemantics from outsiders, arise not from irreconcilable principles about the nature of representation, but rather from the dispensable assumption that perceptual representation is a single relation.

Could we eschew selectionist functions and base a theory of perceptual representations entirely on teleological Cummins-functions? This isn't an advisable course. As is indicated by the mitigating factors I mentioned above, representations based on selectionist functions form a large and important class, incorporating many learned functions as well as functions that are inherited. They deserve to be recognized independently.

This brings me to a second revision of the simple selectionist interpretation of (TS). Consider again the selectionist account of functions that Neander offered:

> It is a/the proper function of an item (*X*) of an organism (*O*) to do that which items of *X*'s type did to contribute to the inclusive fitness of *O*'s ancestors, and which caused the genotype, of which *X* is the phenotypic expression, to be selected by natural selection. (Neander 1991, 174)

On reflection this proposal seems a bit narrow. It perfectly fits the case in which tokens of a perceptual state R have the function of encoding information about property P because (i) earlier tokens of R encoded information about P and (ii) this led to R's being selected. But a token of a state R can also have the function of encoding information about a property P even if no past token of R has ever

encoded information about P. This would be true, for example, of a token of a representation of a shade of color that had never before been experienced by a human being. The token would not owe its informational function to its being a token of a representation-type that was selected itself, but rather to its being a token of a representation-type R that is affiliated with a system of representation-types Σ, where Σ was selected because a great many of its constituent representation-types had a history of encoding useful information about colors. That is, R was not selected on the basis of its own merits, or the merits of its tokens, but rather on the basis of the merits of other representations that belong to the same family as R. It follows that Neander's proposal doesn't tell the whole story about informational functions. It can and should be generalized to allow for representational contents that individual states have as components of representational systems. (Neander acknowledges and endorses this point in a later work (Neander 2017, Chapter 8).)

I conclude this discussion of informational functions by noting that Shea (2018) and Williams (2019) also favor pluralistic accounts of perceptual representation. Moreover, Williams shares my view that a full account of perceptual representation calls for both selectionist functions and Cummins-functions, though his proposal differs considerably in motivation and also in details from mine. Shea and I diverge at several major choice points. Perhaps the most important of these involves the question of what is involved in the acquisition of representations by learning: while I strongly favor accounts of learning that restrict it to biological systems (and certain systems produced by human design), Shea prefers broader, more inclusive accounts that attribute learning to simple mechanical systems that provide no foothold for such biological concepts as need, goal, and reward (or for their counterparts in the realm of design). (See, e.g., Shea 2018, Sections 3.6 and 4.1.)

VI. Two Objections to the Foregoing Picture

The foregoing picture is a version of the view that biologically based functions figure in perceptual representations of all types. Perhaps the most widely discussed of the objections to this view asks us to imagine a creature ("Swampman") who comes into existence as a result of quantum mechanical fluctuations in swamp gas, and who therefore has no biological ancestry, but who is otherwise just like ourselves (Davidson 1987). He is made out of cells, has a skeleton, muscles, and a full set of organs, and is guided by perceptual systems that are isomorphic to ours. It is maintained that Swampman must have perceptual representations, and that he is therefore a counterexample to all biological interpretations of (TS). I discuss this objection in Section V of Chapter 5. If the reader gets that far in the book, it will be apparent that I hold the objection in very low esteem.

As I see it, it is an attempt to use an a priori thought experiment to settle an issue that is ultimately empirical in character. Nothing good can come from efforts of that sort.

Another worry is that representational pluralism could lead to a ridiculous proliferation of representata. If, say, we decide that a visual representation stands for the external property *small, black, and moving*, why not also say that it stands for the corresponding property of packets of light immediately before the frog's eye, the corresponding property of images on the frog's retina, and the corresponding property of activity in the frog's optic nerve? I made a few remarks about a special case of this concern earlier (the worry that *flying insect* is represented), but it requires further consideration.

An important part of the answer is that the causal/mechanical property *small, black, and moving* is instantiated by the same objects as the ecological property *food*. This is not true, for example, of retinal images of small, black, moving objects. Equally, retinal projections of items occupying points in Tsao's quality space are not faces. But these observations lead to the more general question of what the grounds might be for the assumption, common to folk psychology and to perception science, that the objects of perceptual experience are external. Further, what grounds might there be for the assumption that the objects of visual and auditory experience are not just external but distal, in the sense of standing at some remove from the sense organs they activate? I think the way to approach these issues is to consider the criteria that perception scientists use to settle questions of representation. Assuming scientific realism, that should lead us to the correct view. I will conclude this chapter by saying something about those criteria.

A perception scientist approaches the domain of perception with a commitment to certain broad biological principles and to certain explanatory objectives.

Most importantly, the biological principles are theses about the externally directed needs and abilities of the species that is being investigated. To focus on the case of human beings, it is assumed that they have a number of fundamental, externally directed needs—specifically, needs for food, protection against predators, shelter from the elements, social interactions with conspecifics, and opportunities to mate. It is also assumed that humans have goals determined by these needs, together with a broad range of externally directed abilities that enable them to achieve the goals, including principally the ability to identify objects and conditions corresponding to fundamental needs when they are encountered, the ability to navigate environments in pursuit of such items, and the ability to act appropriately when the items are achieved. As Tyler Burge has emphasized, these biological assumptions determine a complex ontology consisting of external objects and properties. They are the objects and properties that figure in fundamental human needs, in human goals, and in the corresponding behavioral capacities. (Burge 2010; Shea (2018, 90) makes a closely related point.)

This brings us to the explanatory objectives of perception science. As Burge has also emphasized, these objectives are focused on the foregoing ontology. More specifically, perception scientists take it to be their primary responsibility to explain how perception supports human interactions with components of the ontology, and they have come to believe that the best way of fulfilling this responsibility is to figure out how the human perceptual systems manage to represent such components. In the case of vision science and the science of audition, the task is to explain how the perceptual system manages to represent *distal* components of the ontology. Accordingly, if there is a choice between two hypotheses about what a given visual or auditory state represents, a distal object X that is a component of the ontology of perception, or a more proximal entity, such as a retinal projection or a packet of light that is reflected by X, the hypothesis that the state represents X is automatically preferred. The literature of perception science amply attests to the truth of these claims.

This explains why perception science is concerned with representations of such things as berries, bears, rocks, caves, trees, and conspecifics, but it must also be explained why it is focused on representations of certain low level properties of such objects. To illustrate, in the case of vision, it is focused on representations of sizes, shapes, colors, textures, velocities, and distances. Why is this? The obvious answer is that these are the properties of berries, bears, and so on that can causally influence the processing mechanisms in the visual system. The mind keeps track of berries, etc. by keeping track of their low level properties.

It might seem that in appealing to the commitments and practices of perception scientists I am advocating some sort of stance-relativism. On such a view, perceptual states do not have representational contents autonomously and categorically, but only relative to the explanatory agenda of some group of human inquirers (Dennett 1987). This is most definitely not what I believe. Given scientific realism, appealing to the antecedent commitments and explanatory interests of perception scientists is just a way of pointing to a domain of objective entities and to a set of factual questions about entities in that domain. Thus, given scientific realism, it is true quite independently of any agenda of inquirers that there is a domain of human biological needs, goals, and actions, and an ontology of distal external objects that is delimited by this domain. Also, quite independently of any agenda of inquirers, there are questions about how humans manage to negotiate this domain. Moreover, the answers to these questions are truths that perception scientists discover, not frameworks that they impose on data. It is, for example, quite definitely a discovery that the rat hippocampus contains cells that have the function of encoding information about a rat's location in an external environment (O'Keefe and Nadel 1978, O'Keefe and Burgess 1996). In general, there is no more reason to think that the findings of perception scientists are agenda- or

perspective-relative than to suppose that this is true of discoveries in biology or physics.

I add that the folk theory of perception seems to incorporate assumptions about human need, goals, and actions like those of perception science, and also to have roughly similar objectives, though of course its explanatory ambitions are much more tightly circumscribed.

VII. Conclusion

The goals of the book are to advance the cause of representationalism, to elevate the traditional discussion of appearance and reality, to provide an account of perceptual phenomenology that is compatible with physicalism, to explain the main relations between perception and high-level cognitive states, and to promote externalist views in epistemology and philosophy of mind. In pursuing these goals I will be focusing on perceptual experience, for that is the locus where the goals intersect. The discussion will not, however, be confined to perceptual experience.

In the present chapter I have been concerned to introduce these themes of later chapters, but I have also maintained that it seems possible to develop representationalism in a way that resolves the disagreements in the current literature. Convergence and progress may be possible. The key to this, I have suggested, is a version of representationalism that incorporates teleosemantics, representational pluralism, selectionist and Cumminsian accounts of functions, and an assumption about the objectives of perception science. I will not be relying on this collection of views throughout the book. Many of my arguments presuppose only a highly abstract and generic form of representationalism. But the collection of views will orient the discussion and in some cases provide motivation for premises.

2
Appearance and Reality I

I. Introduction

It is natural to believe that we are aware of the objective, physical properties of things—that is, properties like physical sizes, shapes, distances, textures, velocities, and so on. This *must* be true, one supposes, because the success or failure of our actions depends on such properties. The fact that we generally succeed in tasks like navigating, throwing, catching, climbing, and grasping shows that we are in some sense aware of objective properties.

What explains this awareness? A number of philosophers hold that *perceptual experience* provides us with a grasp of objective physical properties. To illustrate, this view has been defended by Tyler Burge (2010), and also by various schools of naïve or direct realism, including the so-called disjunctivists (Byrne and Logue 2008). Their view has a strong intuitive appeal. It is natural to think that we perceive objective sizes, shapes, distances, and so on, because it seems that our actions are guided by perceptual experience. If this is true, and it is also true that the success or failure of our actions turns on objective properties, then the fact that our most basic actions are usually successful seems to entail that perceptual experience puts us in touch with such properties.

There is, however, a long tradition in philosophy and psychology of rejecting the idea that perceptual experience presents us with objective properties. People in this tradition have instead embraced the doctrine of *perceptual relativity.* Here is a formulation of this doctrine that captures its scope:

For any situation S, the ways objects perceptually appear to one in S are influenced by a number of factors that are independent of the intrinsic natures of the objects themselves. More specifically, the ways objects perceptually appear to one in S are influenced by factors of the following three kinds: first, environmental conditions like distance from the object, angle of view, and lighting; second, features of peripheral sensory processors such as cone density and the tuning curves of individual cones; and third, internal conditions like attention, adaptation, and pathology.

In one form or another, this doctrine was endorsed by Protagoras, Sextus Empiricus, Hume, and the sense datum theorists of the twentieth century—among many others.

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0002

These authors all believed that the doctrine conflicts with the idea that perceptual experience provides direct access to objective properties. And it is easy to see why they were drawn to this view. If, for example, as the doctrine implies, a physical object of a certain fixed size appears large when seen close at hand and small when seen at a distance, and the object has a range of other apparent sizes at intermediate distances, then it is natural to entertain doubts about the idea that appearances adequately reflect the enduring physical size of the object. Thus, there is an obvious problem with saying that all of the appearances of the object reflect the physical size, given that they are all different, and there is also an obvious problem with picking one of the many different appearances and claiming that it alone is accurate—prima facie, at least, any such choice would be arbitrary. But if the objective sizes of objects aren't adequately reflected in the ways that objects appear to us, how can it be true that perception provides us with immediate, fundamental access to objective sizes? (I am here following the line of thought found in paragraphs 3 and 4 of Hume (2000), Book I, Part IV, Section IV. Hume's reasoning in that passage is explicitly concerned only with secondary qualities, but a companion argument in paragraph 9 of Hume (2007), Section XII, Part I shows that it is meant to apply to size and other primary qualities as well.)

This traditional line of thought has an obvious appeal, especially to those who share the intuition that our perceptual awareness of objects is grounded in the ways that objects appear to us. But I will present additional reasons for the view. To illustrate, as we will find in Section III, contemporary work in vision science shows that perceptual judgments concerning even very basic objective properties, such as distances and sizes, are pervasively distorted, often falling short of accuracy by wide margins. These distortions in our perceptual *judgments* are best explained by the hypothesis that perceptual *experience* is pervasively out of step with objective properties.

It appears, then, that there is tension between perceptual relativity and the idea that objective physical properties are presented directly in perceptual experience. I will be concerned with this tension in the present chapter. I will begin by giving arguments of several different kinds for perceptual relativity, and will then maintain that relativity provides strong reasons for holding that perceptual experience is not concerned with the enduring objective properties of external objects, but rather with *appearance properties*, which I take to be viewpoint-dependent and highly volatile. The next step will be to sharpen our conception of appearance properties. Among other things, this will further challenge the idea that we have a purely experiential grasp of objective properties. Finally, I will propose an alternative account of how objective properties are known. The over-arching idea of this account is that we do not achieve contact with objective reality via perceptual experience alone, but via perceptual experience acting in concert with memory and our cognitive and motoric systems. To illustrate, the account claims that our

grasp of objective reality derives in part from the systems that control action. These motoric systems bring experiential representations into alignment with reality by "interpreting" or calibrating them with respect to objective distances, objective sizes, objective shapes, and so on.

Partly with a view to simplifying the discussion, and partly because more is known about vision than other perceptual modalities, I will focus on vision and visual experience. Also, I will generally focus on visual awareness of spatial properties rather than awareness of such properties as color and velocity.

A terminological note. I have been contrasting the objective physical properties of external objects with their relational, viewpoint-dependent properties, but, as will emerge, there is a sense in which the relational, viewpoint-dependent properties that I have in mind are perfectly objective. They exist independently of the minds and perceptual systems of observers. That is, while they are view-*point* dependent, they are not view-*er* dependent. Nonetheless, they contrast sharply with what I have been calling objective physical properties. The properties I have been calling objective and physical are intrinsic, possessed by objects independently of their relations to different viewpoints.

II. Introspective Grounds for Perceptual Relativity

Consider the following examples:

1. Apparent size: Suppose a car is driving away from you. As it recedes into the distance, it looks smaller and smaller until it finally disappears from view.
2. Apparent shape: Suppose you are looking at a cup from the side. The circular rim at the top of the cup looks elliptical. Or suppose you are looking at parallel railway tracks, or parallel yellow lines marking the edges of a road. As the tracks or lines recede into the distance, they appear to converge.
3. Apparent distance: Suppose there is a sagittal line that begins at your feet and extends outward for a considerable distance. Equal intervals on this line will appear to you to have different lengths, with the closer intervals appearing to be larger. This effect is made vivid when you view broken white lines that indicate a passing zone on a highway. As the distance between you and a line increases, the apparent distance between the endpoints of the line is increasingly compressed. The lines appear shorter and shorter the farther they are away from you.
4. Apparent color: Suppose you are viewing a room that is uniform in color and that has a single light source. The light falls unevenly on different portions of the wall, and as a result, the color of the wall appears to vary across its surface, appearing lighter in some places and darker in others.

5. Apparent velocity: Suppose you are standing at the side of a straight road that extends into the horizon. Looking towards the horizon at cars that are at a considerable distance, it seems that they are moving quite slowly, perhaps as slowly as an ant that is crawling across your floor. But the cars that are whizzing by you seem to be traveling at a much greater speed, even though their speed is of course the same as that of the distant cars.

I hope I have chosen examples that resonate with the reader.

Perhaps it will be useful to elaborate on one of the examples. The fact that objects look smaller as they recede into the distance is known as *size underconstancy*. Now objectivists often allow that perception of objects at great distances shows underconstancy of this sort. After all, who could deny that houses seen through the window of an airplane look smaller than they would if you were standing next to them? It is maintained, however, that we have a good perceptual command of the objective sizes of objects that are closer to us. Sometimes it is thought that this command extends to all objects that are located within so-called personal space, which encompasses an area around the body 2 or 3 meters in diameter. Others have made the much stronger claim that it extends to all objects that are located in action space, which has a diameter of about 30 meters. While different in obvious ways, these views are alike in claiming that there is always an area around an agent in which more or less perfect size constancy is maintained. By the same token, they agree in claiming that there is always an area in which visual experience puts agents in touch with objective sizes.

As against these views, I wish to claim that objects look smaller and smaller as they recede within action space, and even as they recede within personal space. There are several introspectively grounded reasons for thinking this. One is just that careful attention to one's experience shows it to be true. Another, slightly more theoretical reason is based on the observation that awareness of motion depends in large part on changes in apparent size. We are aware of motion away from us even when the moving objects are a few inches in front of our faces. Another reason is that smaller objects can occlude much larger objects when they are at different distances within personal space. A fourth reason has to do with the continuity of decrease in apparent size. Suppose that an object is placed at different distances within personal space from the perceiving agent, and that it is then moved outside of personal space. If there was constancy within personal space but underconstancy outside it, the apparent size of the object would undergo a change for the first time when the object exits personal space, and this change would be noticeable. Thus, suppose that an object X is moving from a point P1 immediately in front of your eyes to a point P2 at the boundary of your personal space. Suppose also that the visual system represents X as having its constant, objective size until it arrives at P2, but thereafter represents X as having

angular sizes. The angular sizes subtended by X during its progress from P1 to P2 will have decreased quite substantially (recall that personal space extends for two to three meters), with the result that there will be quite a falling off from the objective size represented just before P2 to the angular size represented immediately afterward. In my experience, anyway, no such discontinuity can be detected. There is a gradual decrease in apparent size, not constancy followed by a sudden drop off. The same is true of action space.

I suggest that what we are observing here with apparent size also holds for apparent shape, apparent distance, apparent color, and so on. Perception of these properties never achieves full constancy. The failure to achieve constancy may be much less noticeable in some contexts than in others. For example, the way apparent size shrinks as a function of distance is much less noticeable for nearby objects than for objects that are farther away. But considerations like those we just reviewed involving apparent size indicate that in those cases, what's true is just that the shrinkage is too small to attract attention, not that it doesn't exist. *A*ttention isn't *grabbed* by the shrinkage, but the shrinkage is *revealed* by attention that is *voluntarily bestowed*.

III. Experimental Grounds for Relativity

There is a long history of empirical work on questions of appearance and reality. The experiments have by no means spoken with one voice, but I think it is fair to say that experiments done in the last twenty-five years have tended to confirm the relativity thesis.

This is true, for example, of Carl Granrud's very interesting work (2009, 2012) on size constancy in children. More specifically, Granrud found that 5–10-year-old children have a strong tendency to underestimate the sizes of nearby objects (objects at a distance of 6.1 meters) by 10%, and a strong tendency to underestimate the sizes of more distant objects (61 meters) by 16%. Granrud's subjects registered their estimates of size by pointing to one of nine nearby comparison objects. By choosing a comparison object of, say, 35 cm, they indicated their judgment that the distant object was also 35 cm in size, or in the neighborhood of that amount. Figure 2.1 shows a subject making such a judgment.

Given that the subjects' judgments took the form of selecting comparison objects, Granrud's results cannot be ascribed to their not having mastered the skill of giving verbal estimates of size. Further, it is important that Granrud asked his subjects to indicate which of the comparison objects matched the targets *in size*. That is, he asked about objective sizes, not apparent sizes. The same is true of all of the other experiments I will mention here. In all cases, experimenters asked their subjects to make judgments about objective properties, not about appearances.

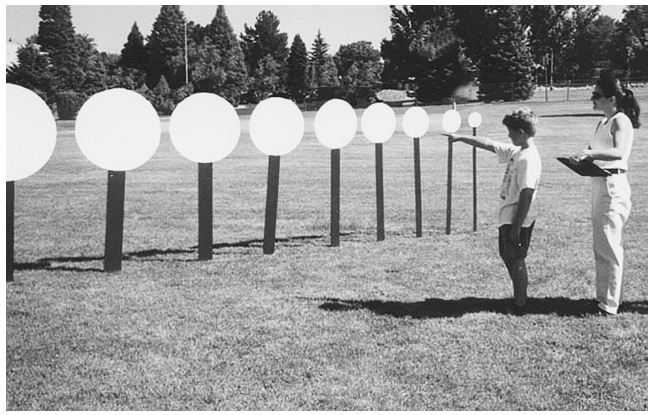

Figure 2.1. A Granrud subject participating in the experiment.
(Thanks to Granrud for providing this image.)

I emphasize this last point because there is a tradition in philosophy of approaching questions about appearance and reality from a linguistic perspective—that is, by doing some sort of semantic analysis of terms like "looks" and "appears." The idea, I think, is that those questions are fundamentally linguistic in character. In my view, nothing could be farther from the truth. Questions about appearance and reality should be seen as questions about the relationship between representational states of the perceiving subjects and the external objects that cause those states. More specifically, they are questions about the representational contents of states of perceiving subjects and the properties of external objects. Empirical inquiries of the sort that I will review here show that there are ways of getting at those questions without considering linguistic issues. One can, for example, find out how a subject is representing the size of a distant object by proceeding as Granrud did—by asking the subject to compare the objective size of a target object with the objective size of a nearby item. Terms like "appear" play no role in such instructions.

To amplify, the topic of interest here is the nature of the properties that are represented at the level of perceptual experience. This is not a linguistic issue. To be sure, it can approached by asking subjects how objects look or appear to them in various circumstances, and when experimenters adopt this form of inquiry, it is crucial to determine how subjects in the experiments understand words like "look" and "appear." Inquiries into the range of meanings that these words have in ordinary language can be relevant here. (Cf. the discussion of Chisholm's distinctions among meanings of "appear" in Chapter 1.) But as Granrud's inquiries illustrate, it isn't necessary to adopt this methodology in exploring the main issue. One can instead proceed by asking subjects to make comparative judgments about objective properties. Moreover, the results of these studies can help to stabilize inquiries that make use of the language of appearances. In evaluating such

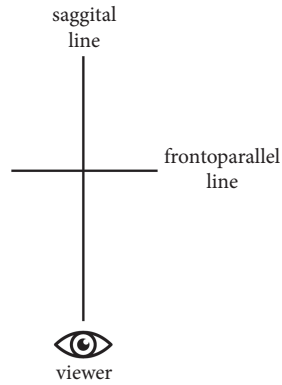

Figure 2.2. Positions of viewer and intersecting lines in Loomis's experiment.

inquiries, special weight should be given to ones that are compatible with results obtained by using the "objective methodology," and in conducting new inquiries, investigators should be careful to disambiguate questions involving terms like "looks" or "appears" in ways that have been found to secure compatibility in past experiments.

I turn now to a famous experiment by Jack Loomis (1996) on the perception of exocentric distances. (Exocentric distance is absolute distance between locations that are determined independently of relations to the subject. It contrasts with egocentric distance, which is distance from the viewing subject.) Loomis showed his subjects pairs of intersecting lines. One member of each pair was sagittal in orientation and the other was frontoparallel to the subject (Figure 2.2).

The subject's task was to adjust the length of the sagittal line until it appeared to be the same length as the frontoparallel line. Loomis found that the physical length of the sagittal line had to be 50%–90% greater than the physical length of the frontoparallel line in order for the lines to appear to the subject to be equal. In other words, there is considerable compression of apparent exocentric distance. This result is of course of a piece with our earlier observation that the stripes marking passing lanes in highways seem to decrease in length with distance. It strengthens the case for that earlier observation, and at the same time shows that it points in the direction of a quite general principle. (For similar results concerning personal space, see Norman et al. 1996.)

I have reported Loomis's result using the term "appear" in the interests of expository economy, and I will continue to allow myself this luxury in reporting other results, but, as noted a moment ago, in all the experiments I will review, the language experimenters actually use to communicate with subjects is free from terms of this sort.

Loomis's result is concerned with perceived exocentric distances. There are also results indicating that perceived egocentric distances are underestimated

significantly. Some of the relevant data involve verbal estimates (Foley, Ribeiro-Filho, and DaSilva, 2004; Loomis and Philbeck, 2008), but others involve non-verbal behavior. Thus, Li, Phillips, and Durgin (2011) arranged subjects in such a way that they were facing an end point, A, of a frontoparallel line running from A to B, and asked the subjects to walk toward A until the distance between them and A was the same as the distance between A and B. Their subjects stopped at distances that were 30% greater than the ones they were trying to match. Like exocentric depth, egocentric distance is highly compressed.

Another example is afforded by Frank Durgin's work on perceiving the slants of hills and ramps (Durgin and Li, 2017). Durgin has shown that the apparent slant of slopes varies with the distance of the observer. Slopes look steeper when judged from a distance than when judged from nearby positions. (Durgin and his colleagues have also obtained several other important results indicating divergences between appearance and reality in the case of slants.)

A result of a different kind due to Durgin is that there is a systematic divergence between appearance and reality in the case of optic flow (Durgin 2009). Pace of optic flow refers to the rate at which visual stimuli succeed one another—where the succession may be due to motion of the subject relative to the environment, motion of the environment relative to the subject, or, as in Durgin's experiment, virtual reality equipment worn by the subject. Durgin equipped his subjects with VR glasses and provided them with a virtual environment in which a hallway with spots on the walls appeared to be moving past them. Durgin asked his subjects to estimate the pace of optic flow both when they were walking and also when they were standing still. He found that their estimates of the objective pace of flow in the first condition were significantly lower than their estimates of the objective pace in the second condition. That is to say, a fixed rate of optic flow (as measured by the duration of VR images) presents a different appearance to subjects who are in motion than to subjects who are standing still, even though the objective optic flow is the same. Here is yet another way in which appearances change while an objective property remains constant.

There is also important experimental work on perspectival shape. We noticed in the previous section that the round rim of a cup looks elliptical when the cup is viewed from the side. It is also true that a circular coin looks elliptical when it is slanted away from the viewer. Or at least, this is how we are inclined to describe our experience when we consider it introspectively. But there is experimental work that confirms these impressions. Older work in this area is summarized in Epstein and Park 1963 and Daoust 2021. More recently, Morales et al. (2020) showed subjects pairs of objects. A pair might consist either of a genuine ellipse and a slanted circle or of a genuine ellipse and circle presented head on. Subjects were asked to respond by indicating which member of a pair was the genuine ellipse. They took longer to answer when the displayed pair consisted of a genuine ellipse and a slanted circle than when the displayed pair consisted of a genuine ellipse and a head-on circle.

The most natural interpretation of this result is that the appearance of the slanted circle was similar to the appearance of the genuine ellipse. (For objections and persuasive replies to those objections, see Burge and Burge (forthcoming) and Morales and Firestone (forthcoming), respectively.)

The examples so far have concerned spatial and kinematic properties. There is also interesting work demonstrating underconstancies in other areas, including color perception and the perception of surface features like glossiness. To illustrate, experiments by Marlow et al. (2012) indicate that perceived glossiness varies with the structure of the light field, the observer's viewing position, and 3D surface geometry.

The studies I have been reviewing are just drops in the sea of experimental work on perceptual appearances. But I have not been cherry picking. Rather, I have chosen studies that are widely regarded as dispositive, due to the excellence of their experimental design and/or their replicability. Moreover, they reflect trends in the larger literature. Thus, for example, in his meta-analysis of the literature on children's assessment of sizes, Mark Wagner (Wagner 2006) arrives at conclusions that cohere nicely with the more recent work done by Granrud. Again, in his meta-analysis of work on perception of sagittal distance, Wagner concludes that 17 of 18 studies found substantial degrees of underconstancy, no matter whether subjects were asked to assess apparent length or objective, physical length. The results of these studies cohere nicely with Loomis's results about the perception of distance. To be sure, there are plenty of divergent data in the empirical literature on appearances; but, as Wagner points out, it is possible that many studies are flawed by ambiguities in their instructions to subjects.

So far I have been considering experiments indicating (i) that external conditions of perception, such as the distance from the observer, can cause appearances to diverge from reality. But this is only part of the story. In addition to claiming that the visual appearances vary with external conditions, perceptual relativity also claims (ii) that appearances vary with features of sense organs, such as the distribution of cone cells and the photopigments that attach to them, and (iii) that appearances vary with internal factors such as attention and adaptation. In addition to the experimental evidence for (i), there is also considerable experimental evidence for (ii) and (iii). I will only be considering questions concerning (i) in the present chapter, but evidence for (ii) and (iii) will be discussed in Chapter 3.

IV. Appearance Properties

It is clear, then, that there are strong reasons for accepting the doctrine of perceptual relativity. Further, as reflection shows, these are at the same time reasons for doubting that the objective physical properties of objects are fixed by the ways that objects appear to us—by the ways we experience them. As Hume and other

historical figures emphasized, appearances are highly perspectival and in constant flux, while objective physical properties are comparatively stable. Moreover, as the experiments we have been reviewing show, appearances routinely cause us to make *substantial* errors about objective physical properties. Hence, if we were to say that experience represents objective properties, we would be saddled with the view that it systematically misrepresents them, often by a wide margin, in a number of different respects. Indeed, we would have to acknowledge that the systems of representation involved in visual experience rarely if ever get things right, and often miss by large margins. This is not a comfortable conclusion. Our best theories of representation (e.g., Millikan (1989a), Dretske (1995), Burge (2010), Neander (2017), Shea (2018)) all speak against the possibility of representational systems that are pervasively erroneous. (There is more detailed criticism of this idea in Section I of Chapter 3.)

But this leaves us with the question, if perceptual experience doesn't represent objective sizes, shapes, distances, and so on, what exactly does it represent? That is to say, what is it that we're perceptually aware of when we look out at the world? The answer that makes most sense to me, and that I will defend here, is that perceptual experience puts us in touch with *appearance properties*, which I take to be relational, viewpoint-dependent properties of external objects.

As will be seen, there are positive arguments for this view that are independent of the forementioned reasons for doubting that perceptual experience represents objective physical properties. These arguments strengthen the case against the latter view.

For the present, I won't be making any concrete claims about the nature of viewpoint-dependent properties. Instead, I'll just cite some examples of viewpoint-dependent properties and say that what I have in mind is something like properties of that type.

My examples are *visual angles*. Visual angles are the angles that external objects subtend with respect to the eye of the viewing subject. Figure 2.3 shows an example. The visual system can uniquely recover the visual angle subtended by an object from the size of the image that the object projects on the retina.

Visual angles are a natural candidate for the role of apparent sizes. This is because, like apparent sizes, they decrease smoothly as a function of distance. Another reason is that they are viewpoint-dependent, not viewer-dependent. They can be defined with respect to points in space—points that may or may not happen to be occupied by lenses of eyes. Saying that size appearances are visual angles has no tendency to imply that size appearances are properties of mental states, or properties of states of the visual system. This is a virtue because, as many people have pointed out, perceptual experience is transparent: it seems to present us with properties of external objects, not with properties of our own internal states. (See, e.g., Harman 1990.) Theories should respect fundamental intuitions to the extent that they can. (There will be more on transparency later.)

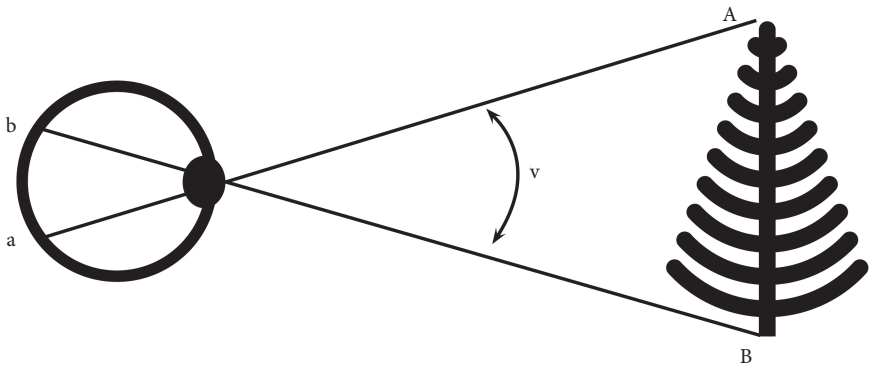

Figure 2.3. The visual angle V.

As I indicated a moment ago, I don't think that visual angles are the best candidates for size appearances. I think the correct story is more complicated than that. But visual angles nicely illustrate the appeal of taking appearance properties to be relational, viewpoint-dependent properties. That is my reason for citing them here.

Now that we have a general idea of what appearance properties might be, I want to make the further claim that appearance properties are the properties we are aware of in visual experience. It might seem otherwise, because we often describe appearances by propositions of the form *x looks F to y*. Propositions of this form suggest that appearances are *modes* of awareness, not *objects* of awareness. The only objects of awareness recognized by such propositions are values of the variable x, and the values of x are objects that appear to us, not appearances of objects. Reflection shows, however, that we also say things that imply awareness of appearances. In discourse and in thought, we acknowledge that we are aware of *how cars look when seen from a distance*, of *how railroad tracks look as they recede towards the horizon*, and of *how tan things look when they are cloaked in shadow*. We are also aware of *how ambulance sirens sound when their sources are moving away from us*, and of *how onions smell when they are held right under our noses.*

In addition to these intuitive considerations about particular cases, there are also general arguments for the claim that we are aware of appearance properties. One argument begins with the premise that we are able to describe appearances and answer questions about them. I can, for example, tell you how this red cloth looks to me at night. The second premise is that these abilities presuppose awareness of appearances. How could I describe the way the red cloth looks if it was outside my ken? It follows trivially that we are aware of appearances. A related argument for the awareness claim is that we seem to be able to attend perceptually to appearances. The chimney of the house next door looks small when it is seen from my window. I can attend to that fact. But perceptual attention is a form of perceptual awareness. Another argument invokes the fact that we are acquainted

with perceptual phenomenology. Perhaps it would be useful to make the structure of this argument fully explicit:

First premise: We are directly aware of perceptual phenomenology. Our grasp of it is not based on descriptions, nor on inferences from something else that we grasp.

Second premise: Perceptual phenomenology consists of appearances. As Jaegwon Kim has written, perceptual qualia are "the ways things *look*, *seem*, and *appear* to conscious observers." (Kim 2006, 225)

Conclusion: We are aware of appearances.

Finally, my favorite argument for the claim is that it is an essential part of a plausible theory of how perception works. This theory will emerge as we proceed. I hope you will agree that it is appealing.

We have arrived at some conclusions that cannot easily be reconciled. On the one hand, it seems that we must somehow grasp objective properties, for the success and failure of our actions depends on the objective layout of our environment, and the actions we undertake to fulfill our basic needs usually succeed. It is also very tempting to hold that it is perceptual experience that provides our must fundamental access to the environment. Thus, it seems that the planning and execution of our actions is guided by perception. But reflection shows that our perceptual experience doesn't represent enduring objective properties, but rather an altogether different class of properties that are viewpoint-dependent and therefore highly volatile. Yikes!

V. Thouless Sizes

With a view to getting a more substantial grasp on the nature of appearance properties, let's consider two hypotheses about the nature of apparent size. It seems that one or the other of these two hypotheses must be correct.

According to the first hypothesis, the apparent size of an object can be identified with the visual angle that the object subtends with respect to the nodal point of the eye. This hypothesis is beautifully simple and straightforward, and, as noted earlier, it implies that the apparent size of an object decreases smoothly as the distance from the observer increases, a prediction that is confirmed by introspection and also by many experimental studies. In view of these virtues, it is no surprise that the hypothesis has found favor with a number of people, including Harman (1990), Huemer (2001), and Noë (2006). The second hypothesis is more complex. Like the first hypothesis, it claims that the apparent size of an object is a relational, viewpoint-dependent property of the object. But, unlike the first hypothesis, it identifies the apparent size of the object with a property that results from applying a certain function *f* to the visual angle that the object subtends.

More specifically, *f* is a function that takes visual angles and various data related to distance as inputs and yields a relational, viewpoint-dependent property as an output. *f* is supposed to be a function that is actually computed by the visual system.

I will argue briefly that the first hypothesis is too simple to be true, and will then urge that the second hypothesis is quite promising.

Contrary to what the first hypothesis claims, visual experience doesn't represent true visual angles, but rather quantities that result from applying certain transformations to visual angles. To see this, consider what is involved when you see two people of approximately the same size on the street in front of you—one at a distance of fifteen feet, and the other twice as far away. Since the first person is only half as far away as the second person, the latter subtends a visual angle that is only half as large as the angle subtended by the former. But the latter doesn't look half as tall as the former. True, the latter looks smaller than the former, but he or she looks much more than half as tall. For another example, extend your left arm twice as far as your right arm, and examine the apparent sizes of your two thumbs. The left thumb will subtend a visual angle that is only half the size of the angle subtended by the right thumb, but the left thumb won't look only half as big as the right thumb. In both of these cases there is a discrepancy in apparent size, but the discrepancy is much smaller than a half. Effects of this kind are pervasive in visual experience: it holds as a rule that the apparent size of an object cannot be predicted on the basis of visual angle alone. (Although this fact tends not to be widely appreciated among contemporary philosophers, it was well known to Descartes. See Descartes 1971, 252.)

Evidently, the reason for this divergence between experienced size and visual angles is that our visual experience of size is shaped by mechanisms that compute constancy functions. That is, there are constancy mechanisms that operate prior to the formation of visual experiences, and that influence the formative process. (There is considerable evidence that constancy mechanisms operate at early stages of visual processing—as early as processing in V1 and V2. See, e.g., MacEvoy and Paradiso 2001, Hurlburt 2003, Murray et al. 2006, Sperandio et al. (2012), and Pooresmaeli et al. 2013.) The second hypothesis allows for these mechanisms, but the first hypothesis does not. Accordingly we must set it aside.

On the second hypothesis, the representations that are constitutive of visual experience represent neither objective sizes nor angular sizes, but rather quantities of a third type that I will call *Thouless sizes*. (I'll explain in a minute why this label is apt.) The Thouless size of an object at a time is computed by mechanisms that operate on visual angles and packages of information about distance. These mechanisms are "partial constancy" mechanisms—that is, they are mechanisms that partially "stabilize" an observer's perceptual experience of an object, making experiences less volatile than the corresponding retinal images, which are as transitory as the visual angles that they register.

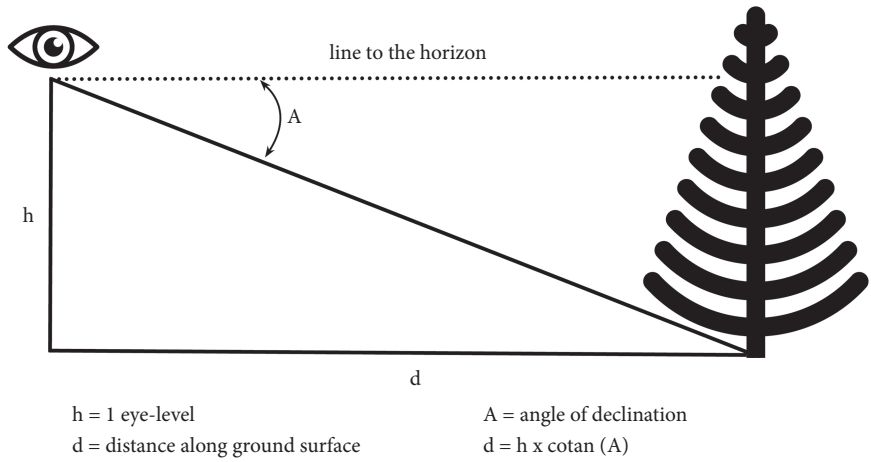

Figure 2.4. Distance from angle of declination.

Let's pause for a moment to reflect on the difference between true constancy mechanisms and mechanisms that only achieve partial constancy. I will describe a mechanism that is capable of producing size constancy.

It is a law of trigonometry that the vertical size, or height, of an object O is a function of the distance from the viewing agent to O and the visual angle that is subtended by O. Now, in principle, the subject has access to both the distance and the visual angle. Thus, the visual angle is given directly to the subject. (More precisely, it is proportional to the retinal image projected by the object, which is clearly directly accessible.) Moreover, the subject is also able to use trigonometry to compute the distance d to O in units determined by his or her eye height (Palmer 1999). Figure 2.4 shows how this works, where O is the base of a tree. As is illustrated here, the distance d to O (the tree) is a computable function of the eye height h of the subject and A, the so-called angle of declination. It follows that the subject has indirect access to the distance to O.

Now the height of O is a fairly simple function of the visual angle that it subtends and the distance d. (Specifically, it is given by the following equation: the height of the object O = 2d x tan (A/2).) The point that we want to notice here is that the output of this function is constant across all changes in distance and visual angle. Because the visual angles subtended by an object decrease proportionally as distances increase, the function produces the same value no matter how the arguments vary. Hence, it is a true constancy function.

The second hypothesis about appearances, which we are considering now, claims that the computations that are responsible for visual experiences of height are different than these trigonometric computations involving true visual angles and true distances in that the outputs show a tendency toward full constancy but in fact never achieve it. There are always underconstancies, in the sense that

apparent sizes always decrease with distance. I call these outputs Thouless sizes, to register the fact that they reflect both sizes and distances, but are not constant, in the way that true objective sizes are.

Let M be the mechanism that takes visual angles and distances as inputs and delivers Thouless sizes as outputs. Further, let F be the function that M computes. Another way to put the second hypothesis about apparent sizes is to say that M produces representations that stand for *Thouless sizes*. The Thouless size of an object x with respect to viewpoint y is F(v, d), where v is the visual angle x subtends with respect to y and d is information pertinent to the distance from x to y. The function F is a "partial-constancy" function, in the sense that its values are more stable than ever-fluctuating retinal images but are nonetheless always characterized by under-constancy.

The doctrine of Thouless sizes provides a useful model for thinking about visual appearances of various other kinds. Generalizing the doctrine, we obtain a picture according to which the contents of visual experiences are the products of functions that take relational, viewpoint-dependent properties as inputs and yield relational, viewpoint-dependent properties as outputs. In all cases, these functions are "partial constancy functions." That is to say, in all cases, the properties they yield as outputs are more "stable" than the inputs, in the sense that they tend to be less susceptible than the inputs to changes in external factors like distance, lighting, and angle of view. They are less volatile than the inputs. But they are nonetheless significantly more variable than the intrinsic, objective properties of external objects: objective properties are more or less constant from moment to moment, at least when intervals are comparatively short, but the functions that the visual system computes tend to yield divergent values across time.

I refer to the outputs of these functions as *Thouless properties*. In so doing, I mean to commemorate R. H. Thouless, an early twentieth-century psychophysicist (Thouless 1931a, 1931b, and 1972) who developed an account of perception and perceptual experience that is similar to the one that I am elaborating here. Thouless wrote that size appearances "are a compromise between the changing size of the peripheral stimulus and the unchanging 'real' size of the object" (Thouless 1931a, 353), and he extended this account of size appearances to appearances of color and shape. More generally, he seems to have thought that all visual appearances are the outputs of "partial constancy" transformations. This is precisely the view of appearances that I'm recommending. (Egon Brunswik advanced similar views at about the same time as Thouless. See, e.g., Brunswik 1956.)

I have been presupposing that the visual system arrives at representations of size by computations involving visual angles and distances. I believe that this is the most plausible view of the matter, but there are alternative ways of thinking about size computations—for example, there is the Gibsonian idea that the visual system performs "horizon ratio" computations. (See, e.g., Palmer 1999, 321.) The picture I have been developing can be adapted to alternative hypotheses of this

kind. Moreover, for an alternative hypothesis to be admissible, it must be possible to fold it into an adaptation of the picture. Any theory of size computations must accommodate the data indicating underconstancies, and must also, therefore, recognize counterparts of Thouless properties as I have defined them here.

VI. More on Thouless Properties

Clearly, if the properties presented in experience are Thouless properties, the story of how we use experience to arrive at awareness of objective properties is going to be complex. If we are aware of objective properties at all, we must achieve that awareness by moving from experience itself to non-experiential representations of some sort. In the next section I will put forward a few ideas about the nature of these other representations and the processes leading to them. Here, however, I would like to address a few additional issues about Thouless properties.

First, I don't wish to claim that the ways in which visual appearances diverge from reality can be explained entirely in terms of mechanisms that compute partial constancies. Experiential representations reflect other influences too. To illustrate, the work by Durgin on perceiving slants that we considered earlier indicates that slant perception is characterized by systematic *over*constancy. Again, experiments by Durgin suggest that the visual system contains a mechanism that multiplies angles of declination by 1.5, creating the impression that such angles are significantly greater than they actually are (Durgin and Li 2017). The mechanism can't be said to compute partial constancies, because the outputs of partial constancy mechanisms aren't uniformly increasing. Different inputs can have the same output, as when a toy car in the palm of your hand has the same apparent size as a large car seen far down the road. In contrast, the Durgin transformation is monotone increasing. It just inflexibly expands angles of declination by multiplying them by a fixed factor. In general, experiential representations show the influence of a variety of factors, as do the appearance properties that they represent.

Second, it might seem that in maintaining that experience represents relational, viewpoint-dependent properties, I am embracing an essentially subjectivist picture of the contents of experience. In fact, however, viewpoint-dependent properties can be perfectly objective, where by this I mean that there is no need to appeal to mental states of the perceiver or processes in the perceiver's visual system to define them. Consider the visual angle V that an object subtends with respect to the point in space where the agent's lens is located. The fact that the object subtends V is a perfectly objective fact about the object. The fact is completely independent of the subject's mental states and visual processing. Indeed, it would have obtained even if the observer had never existed. Now let f be a function of visual angles that is actually computed by the human visual system. Clearly, the

fact that *f* yields a certain value when it is applied to V is another perfectly objective fact, just as all facts about functions and their values are perfectly objective. The "subjectivity" of a computational process has no tendency to show that facts involving the function computed by the process are subjective. (My pocket calculator can compute the addition function, but that has no tendency to show that facts involving the addition function are calculator-dependent.)

Third, I am maintaining that appearance properties are Thouless properties, and that Thouless properties are relational, viewpoint-dependent properties of external objects. Why not say instead that Thouless properties are properties of the retinal images that encode information about visual angles and distances, or properties of higher level states of the visual system? That is, instead of claiming that visual mechanisms compute functions of viewpoint-dependent properties of external objects, why not just claim that they compute functions of properties of retinal images, or functions of images in the thalamus or V1? What is the rationale for invoking external objects in explaining appearances?

There are two answers to these questions. (1) As we observed earlier, we all have strong intuitions indicating that perceptual experience is transparent: it seems to us that we are aware of properties of objects that are external to our bodies, as opposed to properties of parts of our visual systems. Consider, for example, the object you happen to be seeing now. I'm sure you will agree that you have a very strong impression that the object is in a three-dimensional space and at some distance from you within that space. Experiencing the object isn't at all like what happens when you are aware of internal phenomena like visual images, conscious thoughts, or pains in the cornea. Nor is hearing an object at all like having a song run through your head. It would be highly counterintuitive to maintain otherwise. Your experience is *as of* externality. This is a reason for thinking that the objects of visual awareness are external to the perceiving subject, for it is very hard to see how the fact that our experiences are as of externality could be explained by internalist or peripheralist views of the objects of perceptual awareness. Far better to explain them in terms of externally oriented representational contents.

(2) As Millikan (1989a), Dretske (1995), Burge (2010), Neander (2017), and Shea (2018) have all emphasized, the contents of visual representations are shaped by our biological needs.

To elaborate: Our biological needs are focused on external objects with locations in physical space—predators, prey, fruit, obstacles, safe havens, and so on. Accordingly, since vision serves our biological needs, we should suppose that vision represents external objects and also properties of external objects—that is, objects that occupy the space in which we must act and that determine the success or failure of our actions. Now it might seem that the foregoing reflections indicate that, contrary to what might otherwise be thought, vision doesn't serve our biological needs. After all, it presents us with Thouless sizes, Thouless

distances, and Thouless shapes rather than their objective counterparts. But there are grounds for thinking that it is after all hugely beneficial to represent Thouless properties. Consider what it would be like if our experience was characterized by full size constancy. All of the antelope scattered across the savannah would have the same apparent size. This would deprive the conscious mind of one of the main cues that it uses to compute distance, with the result that it would be more difficult to figure out which antelope one would have the best chance of hitting with a spear. (There are a number of pictorial distance cues in addition to apparent size, as we will have occasion to observe later on, but apparent sizes is one of the major ones.) Or, to update the point, all of the traffic lights on Broadway would have the same apparent size, with the result that it would be more difficult to figure out which one to obey. For a different sort of example, consider the fact our ability to the perceive the motion of objects away from us depends on the fact that apparent size decreases with distance. If there was perfect size constancy, we would lack this important motion cue. It is also relevant that a great deal of recent work, mainly due to Marisa Carrasco, indicates that attention magnifies apparent contrast. (See, e.g., Carrasco et al. 2004 and Carrasco 2011.) Compare worlds in which this is true with possible worlds in which it doesn't hold. In the latter worlds, visual experience would be much less useful than it actually is in searching for patterns that involve contrast, and in classifying them. On the whole, it seems that a Panglossian reading of the fact that experience presents us with appearances rather than objective properties seems appropriate. But the original point stands: they need to be appearances of the objects that occupy the physical space in which we act and on which the success or failure of our undertakings depends. (This counters an argument against appearance properties in Lande 2018.)

Fourth, the view that conscious visual representations stand for Thouless properties allows us to avoid the unappealing view that perceptual experience is pervasively erroneous, and generally erroneous by a very wide margin. On the Thouless view, perceptual representations hit their targets. Neither of the main alternatives—the view that perceptual representations stand for objective properties, and the view that they stand for simple viewpoint-dependent properties like visual angles—possesses this desideratum. Instead, they commit us to saying that perceptual experience is a realm of illusion. None of our best theories of representation (e.g., Millikan (1989a), Dretske (1995), Burge (2010), Neander (2017), Shea (2018)) is compatible with such a claim.

Fifth, it is sometimes maintained that we are not aware of appearance properties in the ordinary course of events (Palmer 1999). Rather, we become aware of them in contexts that are defined by special interests, such as an interest in drawing according to the principles of linear perspective, or in contexts that are defined by special perceptual circumstances, such as viewing earth-bound objects from the window of a plane. In these special situations our attention is drawn to appearances, but it is only through this special form of attentive awareness that appearances are ever consciously registered.

Now it is quite true that attention can influence appearances. As we have noticed, this has been established by Marisa Carrasco (Carrasco et al. 2004, Carrasco 2011). It would be a mistake, however, to think that awareness of Thoulesss properties always depends on special activations of attention, and that the norm is awareness of objective properties. This can be seen in a number of ways, but perhaps the most immediately convincing demonstration involves performing an operation on one's own phenomenology. Glance around the environment in which you are located, and then commit to attending just to the appearances that the objects in the environment present. (That is, switch to the stance you would have if you were concerned to draw objects in the environment according to the principles of linear perspective.) There will be a phenomenological change when you activate attention in this way, but I predict that you won't regard it as a major change. You will not be tempted to describe it as a change from viewing properties of one kind to viewing properties of an altogether different kind. Rather, it will just seem to you that the properties you were formerly aware of are now more pronounced and in somewhat sharper focus.

The objection based on attention also suffers from other problems. It simply isn't true that we perform a special act of attention whenever we look at houses on the ground through the window of a plane. Houses seen from above look small even when one is an experienced flyer, and further, even when one is engaged in thought about an altogether different topic while looking through the plane's window. Another consideration has to do with the perception of distance. As is well known, the apparent size of distant objects is one of the strongest cues that the visual system makes use of in representing distances. For example, when you are aware that the cars ahead of you on the road are at different distances, this is largely because the apparent sizes of the farther objects are smaller. But no special act of attention is needed to perceive the different distances of the cars. We perceive their different distances automatically, without special intention or effort. It follows that no special act of attention is needed to produce awareness of apparent sizes.

VII. Objective Properties

We have found that there are considerable discrepancies between the properties that are represented in visual experience and the objective correlates of those properties. Given all these discrepancies, it is clear that, contrary to the intuitive picture with which we started, we don't have a purely experiential grasp of objective properties. But if this is so, how can it be true that we are as successful as we are in such experientially guided actions as navigating, throwing, catching, and climbing? Prima facie, at least, it seems that success in such activities requires a grasp of objective properties.

The problem, then, is that of explaining our ability to flourish in the world given that the objects of experiential awareness are Thouless properties. In a moment I will offer a solution that consists of five claims, but first, I need to address a thought stemming from the discussion of representational pluralism in Chapter 1. The thought is that it could be true that visual experiences represent Thouless properties and *also true* that they represent objective sizes, shapes, and so on. Developing this thought, it might be said that Thouless properties account for the phenomenology of perceptual experiences, and might therefore be said to be their most immediate objects, but that there are probabilistic relations between Thouless properties and objective properties that enable the latter to count as secondary or indirect objects. Now, as the reader knows, I am a fan of representational pluralism: I believe it to be a useful tool for resolving certain disagreements about representata that have roiled the literature. I do not think, however, that it is relevant here. This is because perceptual experience does not by itself give us the kind of access to objective properties that could be useful in guiding navigation and judgment. Yes, it is true that Thouless sizes, shapes, and colors are linked probabilistically to objective sizes, shapes, and colors, to some degree anyway, but our experience of Thouless properties does not articulate the latter features, and we are therefore unable to exploit these probabilistic relations in our commerce with reality. For example, my highly perspectival experience of a distant airplane does not by itself allow me to determine whether the plane is large enough for me to fit on board. Perhaps I could only ride astride it, or play with it as a toy. What we are seeking in the present section is representations that give us sufficient command of external reality to be able to act on it effectively.

Are there such representations? What follows is a five-part answer to this question.

1. In many cases our grasp of objective properties is partly cognitive, resulting from cognitive strategies that take representations of Thouless properties as inputs and yield judgments concerning objective properties as outputs. It is obvious that such strategies are at work when perceptual processing results in more or less precise *quantitative* perceptual judgments, such as the judgment that a distant object is about five feet in height. That judgment clearly presupposes cognitive mastery of a system of measurement, and also an ability to apply that system to objects at a range of distances. But cognitive strategies may also be at work in much simpler cases. In a minute I will give an example involving children who are too young to have acquired the ability to apply precise systems of measurement to distant objects.

A cognitive strategy for estimating values of an objective property will make use of information about Thouless properties, but it must make use of additional information if it is to arrive at estimates that do a better job than Thouless properties of capturing the relevant objective values. Consider, for example, a strategy for estimating objective sizes. If its only inputs were Thouless sizes, there would be no reason to expect that its outputs would have any bearing at all on objective

sizes, let alone a reason to expect that its outputs would reflect differences in objective sizes with any degree of reliability. Fortunately, there is good reason to believe that cognitive strategies have access to additional information. Thus, to continue with the example of size, there are many sources of information that are relevant to this magnitude, including all of the following sources of information about distance: accommodation, aerial perspective, angle of declination, binocular disparity, convergence of parallels, edge interpretation, gain of visible mereological structure (as a function of distance), loss of visible mereological structure, motion parallax, occlusion, ocular convergence, position relative to the horizon, shading and shadows, texture accretion, texture deletion, and texture gradients. (For discussion see Palmer 1999, 203–49.) Now it is no doubt true that a number of these sources are exploited by the visual processing that precedes conscious experience, and might not be available to cognitive strategies that operate post-experientially. To illustrate, this is probably true of accommodation, binocular disparity, and binocular convergence, for there is no explicit conscious representation of these magnitudes. But other magnitudes *are* explicitly registered at the level of experience, including aerial perspective, convergence of parallels, gain and loss of visible mereological structure, occlusion, position relative to the horizon, shading and shadows, and texture gradients. In short, the conscious experience of a subject registers a number of variables relevant to estimating distance. It follows that post-experiential cognitive strategies have a lot to work with as they go about their job of estimating objective sizes. The strategies will no doubt fall short of delivering fully accurate results, but they will significantly improve the agent's prospects of successful action.

An example of a cognitive strategy is a procedure that Carl Granrud calls the "distance compensation strategy." As noted earlier, Granrud showed that the size judgments of many 5–10-year-old children are characterized by substantial underconstancies. Granrud also found, however, that some of the children in this group are able to make correct assessments of objective sizes by making use of a conscious strategy—a strategy based on the principle that apparent size tends to decrease with distance. Granrud determined whether children were using this strategy by giving them what he calls the "size–distance knowledge test." The questions on the test include "If I put this way far over there [pointing out a window across the street], will it really and truly be little, or would it just look little?" and "If I put this right up close to your eyes, will it really and truly be big, or will it just look big?" Children who earned high scores on the test were classified as "high knowledge" subjects. Granrud found that high knowledge subjects made much more accurate assessments of sizes than their low knowledge counterparts, and when asked about the reasons for their assessments, their answers showed that they were using a conscious strategy that exploited their additional knowledge. Unsurprisingly, the older the children were, the more likely it was that they would earn high knowledge status by doing well on the test.

It is pretty clear that the distance computation strategy is a high-level, cognitive method. It isn't purely perceptual. Thus, there is a strong correlation between a child's having the ability to make accurate size estimates and the child's having metacognitive knowledge to the effect that he or she is using the strategy. Presumably it wouldn't be possible to have such knowledge if applying the strategy was a purely visual process. Normally visual processing isn't accessible to consciousness.

In adult life we may not be consciously aware of applying the distance compensation strategy when we estimate objective sizes. It may seem to us that our estimates are based entirely on perceptual experience. But Granrud's findings strongly suggest that if we are not aware of following the strategy, this is only because it has become second nature. We no longer need to represent it explicitly in order to apply it, just as an experienced chess player no longer has to explicitly represent the constraints on admissible moves in order to conform to them.

I suggest that these observations about objective size generalize to other dimensions of objectivity. Our success in representing objective slants, shapes, distances, and so on is due in part to our having learned cognitive strategies that enable us to compensate for the divergence between visual experience and reality. This generalization receives experimental support from the work of Durgin and Li, who have found that skiers, hikers, and engineers are better able to assess distances and slants than people who lack their familiarity with outdoor landscapes (Durgin and Li 2017). Skiers etc. presumably learn how to calibrate visual experiences in terms of motor programs, but they also acquire cognitive strategies, as is shown by the fact that their *judgments* of environmental magnitudes are more accurate than those of subjects who lack their training.

2. All visual representation at the level of experience is perspectival, but it is plausible that there are post-experiential representations that abstract away from perspectives but are nonetheless perceptual or quasi-perceptual in nature. Alternatively, we can say that very likely there are representations that are post-experiential but sub-conceptual. This point is often made by reference to faces. We are able to recognize faces when they are presented from many angles, even when the angles are ones that we have rarely encountered. How this is accomplished is known in some quarters as the *problem of visual invariance* and in others as the *problem of appearance variability*. The problem is widely studied and has been the subject of many theories. (There are reviews in Logothetis and Scheinberg 1996 and Frisby and Stone 2010.) Explaining how we represent physical properties such as objective sizes and shapes is a special case of it. It is not known what form such representations take, but there are plausible models of what they might be like. (See, e.g., Biederman 1987, Bülthoff and Edelman 1992, Tarr et al. 1998, Frisby and Stone 2010.) I will describe two such models here. One assimilates post-experiential representations of objective properties to

representations of faces, and the other explains them in terms of a theory of classification known variously as the *exemplar theory* and the *view-based recognition theory*.

It is now known that there are specific cells that respond to faces when presented from many angles (Quian Quiroga et al. 2005). These are sometimes known as *grandmother cells*, as a gesture towards the idea that there is a single cell that fires wherever one sees one's grandmother, regardless of perspective. There is also evidence that there are single cells that respond to landmarks (such as the Sydney Opera House) seen from multiple perspectives (Quian Quiroga et al. 2005). In view of these discoveries, it is natural to consider the possibility that there are also cells that play a role like the role of "grandmother cells," but with respect to objective physical properties like objective size and objective shape. To illustrate, there might be a certain "objective shape cell" that would fire whenever a torus-shaped object was presented, regardless of angle of view. (Or at least, it might be activated by tori presented from a significant range of angles of view.) If so, there might be two types of representation of torus-shaped objects in the visual system—the multiple perspectival representations that are involved in conscious visual experience, and a single post-experiential representation that could be activated by a range of different perspectival representations. A dual content model of visual processing of this sort would of course go some distance toward explaining our grasp of extramental reality.

According to the exemplar theory of classification, many concepts are associated with procedures for recognizing members of their extensions that are based on memories of past encounters with such members. Consider, for example, the concept of a sheep. The exemplar theory claims that the procedure for recognizing sheep has two components: first, a set of memories of particular sheep that the relevant subject has encountered in the past; and second, a procedure for calculating similarities between the members of this set and current perceptual experiences. When the degree of similarity is sufficiently high, the object represented by the current experience is classified as a sheep. There is a rich literature discussing principles that might determine ways of grouping memories into categories and methods of computing similarities between current percepts and these categories. Instead of attempting to describe this literature, I will refer the reader to the comprehensive summaries in Murphy 2002 and Machery 2009, and to the discussion of Hough transforms in Frisby and Stone 2010. But I should add here that the exemplar-based recognition procedure associated with a concept need not be all there is to the concept. Indeed, in typical cases, the concept will also figure in principles and inferences that arguably are partially constitutive of it. Thus, it is plausible that the concept of a sheep is governed by the principle that membership in the category of sheep depends in part on ancestry (Keil 1989, Gelman 2003).

Although the exemplar theory has usually been presented as a theory of concepts and conceptual classification, it can be transposed to the perceptual domain (Palmer 1999, 452, Frisby and Stone 2010, 191–202), and developed into a hypothesis about how objective properties like intrinsic sizes and intrinsic shapes are represented post-experientially in perceptual systems. According to the hypothesis I have in mind, there are exemplar-based procedures for co-classifying or grouping together perceptual experiences that encode perspectival appearances of objective properties. Thus, for example, according to the hypothesis, each sufficiently mature agent has a procedure for recognizing objective rectangularity that is based on a set of perceptual memories of rectangular objects seen from various perspectives—that is, a set of memories of *perspectival appearances* of rectangular objects. Although most of the memories in the set represent appearances that are at least slightly trapezoidal in nature, there are sufficient similarities among them for them to be grouped together into a class that can serve as the basis for a procedure for recognizing instances of the property *being objectively rectangular.* (For discussion of how similarity among remembered exemplars might be determined, see Koenderink and van Doorn 1977.) The recognition procedure consists in computing similarities between agents' memories of perspectival experiences of rectangular objects and their current perspectival experience. Presented objects are classified as rectangular when the level of similarity is sufficiently high.

The exemplar theory of concepts is supported by extensive experimental work (Murphy 2002 and Machery 2009), and there is also evidence for the adaptation of it to the perceptual domain that I have just sketched (Bülthoff and Edelman 1992, Tarr and Bülthoff 1995, Tarr et al. 1998, Frisby and Stone 2010). But it should be emphasized that exemplar theories are controversial. The present hypothesis is just one of a number of possible ways of approaching the problem of visual invariance. Thus, for example, instead of trying to explain recognition of a particular shape in terms of multiple highly perspectival exemplars, *geon theory* attempts to explain recognition of the shape in terms of a single, non-perspectival, volumetric representation (a geon). According to the theory, each such representation is built up from a stock of primitive volumetric representations that can be assembled by applying a small set of principles of construction. To illustrate, a funnel-shaped object might be represented by a geon constructed by adjoining a representation of a tube to a representation of a cone. (For elaboration, see Biederman 1987, Palmer 1999, and Frisby and Stone 2010. In geon theory, talk of volumetric representations carries no commitment to there being physical shapes in the brain, though it is allowed that the representations might bear similarity relations to each other that are in some abstract sense isomorphic to similarity relations between the volumetric structures of external objects. See the discussion of analog representation in Chapter 8.)

Even though there is no settled account of how objective shapes, sizes, colors, and so on are represented, it seems quite likely that the mind employs post-experiential but sub-conceptual systems of *some* kind for recognizing objective properties. After all, we know that there are powerful procedures for perceptual learning, and that these procedures enable agents to acquire general categories for classifying presented objects (Connolly 2019, Ester et al. 2020). It would be surprising if these procedures did not provide ways of assimilating perspectival experiences of objects to general categories corresponding to objective features.

I should add a caveat concerning the idea that the grandmother cells are perceptual mechanisms, and also a caveat for the counterpart idea concerning the exemplar-based classification. There are reasons for thinking that grandmother cells lie somewhat outside the visual system—specifically in the medial temporal lobe (Chang and Tsao 2017, Quian Quiroga 2017, Tsao 2019). If this is true, and it is also true that there are "grandmother" cells that are charged with tracking objective properties, the latter cells might not be fully perceptual. Similarly, it may be that the exemplar-memories on which exemplar-based classification is based are located somewhere outside the visual system, perhaps in the hippocampus. If so, exemplar-based systems for recognizing objective properties aren't fully perceptual in nature. These points are important. They suggest that the mechanisms I have been describing are best seen as quasi-perceptual as opposed to fully perceptual. Even when this is acknowledged, however, it is clear that the mechanisms have a strongly perceptual orientation. Thus, as usually understood, grandmother cells are activated directly by perspectival perceptual experiences, without mediation by concepts or beliefs. Moreover, exemplar-based classification systems constitutively involve memories of past perceptions.

I add that there is also a third reason for thinking that "quasi-perceptual" might be a better term than "strictly perceptual" for the post-experiential representations I've been considering—specifically, evidence that activity in the prefrontal cortex is involved in perceptual classification (Freedman et al. 2001, Freedman et al. 2003). But the story here is complex: there is also evidence that perceptual learning can involve structural changes in the cortical areas that support early vision (Connolly 2019, Dosher and Lu 2020). It will be some time until we can make a reasonable determination as to how much of what we think of as perceptual categorization is genuinely perceptual. All that is clear at present is that perceptual states and processes account for a considerable portion of it.

3. I will argue now that there are probably ways of representing objective properties that combine experiential representations with representations that are essentially motoric in nature.

Try throwing a stone at a tree. Chances are you will miss if you haven't engaged in that activity for a while. Practice will improve your success rate—considerably if you keep at it for a while—but your improvement won't be due to your having

learned how to make judgments of any kind, especially not quantitative judgments. Instead you will have acquired a general visuo-motor skill—that is, you will have learned how to interpret or calibrate perceptual representations in terms of motor programs. There are many similar examples—learning how to shoot a basketball, learning how to hit a target with an arrow, learning how to play golf, etc.

In a wide range of cases, then, our grasp of objective properties is not the result of vision alone, but rather the result of vision acting in concert with the systems that govern action. Notice that it doesn't matter if visual representations fall short of registering objective magnitudes, as long as what they do represent can be calibrated to objective magnitudes by motor programs. There is good reason to think that we actually possess such programs, having acquired them by long periods of learning how to "interpret" visual representations so as to achieve success in enterprises like the ones we just noticed.

To restate this hypothesis, what I am proposing is a version of the idea that representations of objective properties can take the form of pairs consisting of representations of appearance properties and motor programs.

4. Whatever may be the case with perception of distant objects, we have a pretty firm visual grasp of the objective properties of items in personal space. Accordingly, reducing the distance between objects and ourselves is a very important strategy for achieving contact with objective properties. As we all know, we can perform extremely fine-grained operations on items in personal space—operations like grasping, pinching, molding, and chipping.

A caveat. While acknowledging the value of this strategy, I think we must resist the temptation to explain its success by supposing that we have fully accurate visual representations of the objective properties of objects in personal space. As the reader may recall, I argued earlier, on introspective grounds, that there are significant appearance/reality contrasts even within personal space. This view is also strongly supported by scientific work, including a recent series of innovative experiments by Fulvio Domini and his colleagues (Foley 1980, Johnston 1990, Tittle et al. 1995, Bozzacchi and Domini 2015, Bozzacchi et al. 2016, Campagnoli et al. 2017). Using experimental paradigms involving reaching, grasping, and pinching, Domini and his associates have found that there are significant departures from accuracy in representing object-distance and object-depth in personal space. "Grasps directed at an object specified by binocular information usually end at the wrong distance with an incorrect final grip aperture" (Bozzacchi et al. 2016, 255). Apparently, the visual system doesn't possess mechanisms that can accurately compute these quantities from available data, even though it is possible to imagine evolutionary developments in the visual system that would have provided such mechanisms. The actual mechanisms are suboptimal.

How then is the success of actions with nearby targets to be explained? According to Domini and his colleagues, instead of representing magnitudes like distance accurately at the outset of an action, we fine-tune the action as it reaches

completion, relying on visual and haptic feedback concerning the changing spatial relationship between our hand and the targeted object. Success is due to "late and sudden corrections" of such behaviors as reaching and grasping. After a certain point, "online control is continuously engaged until object contact" (Campagnoli et al. 2017, 21). This is indicated by the fact that success is very substantially diminished in experimental contexts in which visual and haptic feedback is deliberately blocked (Bozzacchi et al. 2016).

5. Our grasp of objective properties is often much more modest than we take it to be. In many cases, we lack an adequate perceptual grasp of objective properties, and we also lack a cognitive strategy or a motor program that can make up for the limitations of purely perceptual processing. An example is our perception of the sizes of objects that are suspended above eye level. It is natural for us to feel that we have a reasonably good experiential grasp of those magnitudes. Consider the case of traffic lights. When asked about the size of traffic lights, most people say that they are about two feet in height. And they are fairly confident in their answers. But the actual size of the lights is nearly twice that amount. Another example of unnoticed systematic misrepresentation involves the lengths of the dashed lines on highways which indicate that passing is permissible. When asked, people tend to give two feet as the length. The actual length is ten feet. Evidently, even though we are very close to dashed lines while in a car, something about the perspective prevents us from arriving at a reasonable visual estimate. It also seems that we have no cognitive strategy for correcting this mistake. We are shocked when we are told the answer.

Here then are five proposals about our grasp of objective magnitudes that seem to go some way towards making that grasp more intelligible, in part by deflating our conception of it a bit. Assuming this is true, the proposals provide some relief from the paradox with which we started, according to which perceptual relativity and successful action pull with great strength in opposite directions.

VIII. Conclusion

We have found powerful reasons for thinking that that there is a substantial divergence between appearance and reality. More specifically, we have found these reasons support the idea that perceptual experience presents us with Thouless properties rather than their objective counterparts. This undercuts naïve or direct realism. We have also found that there are various factors that partially compensate for the limitations of experiential awareness.

3
Appearance and Reality II

I. Introduction

In this chapter I will consider five topics that are closely related to the themes of the preceding chapter but are not tightly integrated with them. One has to do with hypotheses concerning the nature of perceptual appearances that are alternatives to the Thouless property hypothesis. To illustrate, these alternatives include the hypothesis that appearances are objective properties of external objects that are systematically misrepresented and the hypothesis that appearances are properties of sense data. The second topic is the fact that the appearances of objects are influenced by various features of peripheral sense organs, such as the retina, in addition to being influenced by external factors like distance and lighting. The third topic is the fact that appearances are also influenced by varying internal factors such as attention, adaptation, and the ability of perceptual processors to function properly. The fourth topic derives from the popular idea that the properties represented in experience must play a causal role in generating experiences. Thouless properties seem not to play such a role. Finally, I will inquire into the forms of experience associated with perceptual modalities other than vision, trying to determine whether, and if so to what extent, the picture of visual experience developed in Chapter 2 can be generalized so as to apply to these other areas.

II. Alternative Accounts of Visual Appearances

Up to now we have been principally concerned with just one explanation of the facts of perceptual relativity—the hypothesis that perceptual experience represents viewpoint-dependent, relational properties of external objects. But there are of course a number of other ways of accounting for appearances and perceptual relativity. In this section I will consider a representative sample of these alternative theories. I will begin with attempts to explain the facts of relativity in terms of properties that are external to the mind, and will then turn to consider a couple of internalist explanations.

The first externalist explanation plays up the importance of error and illusion in perceptual experience. Thus, for example, it might be claimed that when a distant mountain range looks small to a viewer, this is because the viewer is

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0003

misrepresenting the size of the range. The visual system of the viewer lacks sufficient information about the distance of the range to be able to correct for the minuscule size of the image that the range is projecting on the viewer's retina. Accordingly, it attributes the wrong objective size to the range: it represents the range as objectively small. It might be urged, I suppose, that all of the facts cited in support of perceptual relativity can be explained in this way—that is, by claiming that phenomenology is determined by objective properties, but maintaining that the visual system generally attributes objective properties to perceived objects that they do not really have.

The second externalist strategy, which has been recommended by Frank Jackson (Jackson 1977, 103) and Alex Byrne (Byrne 2016), consists in urging that appearances can be analyzed as conjunctions of objective properties. On this view, if a distant SUV looks small to an observer, it is because it has the conjunctive property *being large and far away*. That is, the property *looking small* just *is* this conjunctive property. Equally, if a tan wall looks dark brown, it is because it has the conjunctive property *being tan and in shadow*, and if a round coin looks elliptical, it is because it has the conjunctive property *being round and at an angle.*

The third strategy tries to accommodate relativity by maintaining that awareness is a three-place relation (Campbell, 2009). The first term of the relation is a conscious subject, the second is an external object together with some of its objective, intrinsic properties, and the third is a context, where a context is a package of factors that includes a specific distance, a specific angle of view, a specific degree of illumination, a specific relative velocity, and various other items that play a role in shaping experience. The underlying idea is that it is possible to explain the ways in which appearances and phenomenology are influenced by relativity in terms of changes in the third term of the awareness relation. Relativity doesn't affect *what* we are aware of, but rather only *how* we are aware of it.

Although these strategies have a certain initial appeal, reflection shows that none of them is ultimately successful.

There are several problems with the first strategy. One is that it commits us to saying that visual experience *pervasively* misrepresents objective properties of external objects, sometimes by wide margins. This is a problem because it is quite difficult to explain pervasive representational error, especially when it is substantial. It is agreed on all sides that representation is closely related to information. A standard view is that if a state X of the visual system represents an item Y, then the visual system must have used X to encode information about Y at some point in the past. A related view is that the information must have been beneficial in some way. There can be exceptions to these principles, particularly in cases in which discrepancies between represented properties and objective properties are modest, but how are the principles being respected if it is claimed that huge mountain ranges are often represented as objectively tiny, and light tan surfaces are often represented as objectively dark brown?

A related, but more concrete problem with the first strategy is that it is unable to account for the role that perceptual representations play in planning. Suppose for example that my visual system represents a distant mountain range as objectively small. If I took this representation at face value, I might form a plan to hop over the mountain when I arrived at its base. By the same token, I might think it unnecessary to pack enough food for a long journey, and I might decide to leave my hiking boots and warm clothing at home. These decisions could prove disastrous when I arrive at the range and find that it cannot be traversed in less than five days. Since I don't usually wind up in such predicaments, it must be the case that I use my visual representation of the range as a platform for drawing inferences about its actual size. But how would that work? By hypothesis, my visual system represents the range as objectively tiny. If I infer from this representation that the range has the property of being objectively large, I must be relying on an inferential engine that yields outputs that conflict with the inputs. What could such inference look like? This seems to be the only proposal worth considering:

Content of perceptual experience: There's a tiny mountain range ahead.

Inference from perceptual experience: There's a mountain range ahead that looks tiny.

Inference from the first inference: When depth cues indicate that an item is quite distant (as they do in the present case), and the item looks tiny, then the item is actually quite large.

The trouble with this proposal is that it is weirdly Byzantine to suppose that we directly infer judgments about appearances from perceptual experiences that represent objective properties, and then go on to infer new representations of quite different objective properties. It is much more natural to suppose that perceptual inferences move directly from perceptual representations of appearances to judgments concerning objective properties. More specifically, it is much more natural to suppose that we infer objective perceptual judgments by one or more of the processes described in Section VII of the previous chapter. Those processes do not take representations of objective properties as inputs, but rather representations of appearance properties. Only the outputs are concerned with objective properties. Accordingly, instead of zigzagging, the transitions between inputs and outputs are smooth. (For additional criticism of this approach see pages 76 and 83.)

Turning now to the second externalist strategy, reflection shows that it fails to capture important commonalities of experience. Consider two situations, one in which I hold a toy SUV in my hand and another in which I see a real SUV in the distance. In both cases, it is true to say that an item looks small to me, where "looks small" is being used in its phenomenological sense, but we clearly cannot say that both the toy and the real vehicle have the conjunctive property *being large and far away.* To be sure, it is true that the real vehicle

possesses that property, but it cannot be that it looks small in virtue of possessing it—not if *looking small* is a property it shares with the toy. Now consider a situation in which you are viewing a portion of a tan wall that is cloaked in shadow, and compare it to a situation in which you are looking at a wall that actually is dark brown. Again, there is a commonality that we can express by saying that both walls look dark brown to you. If we adopted the second strategy for analyzing appearances, we would have to say that they look dark brown in virtue of having the property *being tan and in shadow.* The first wall has that property, but not the second.

According to the third externalist strategy, when one is enjoying a visual experience, one is visually aware of one or more objective, intrinsic properties of an external object. Now these properties are relatively stable, while experience of them is in constant flux. Appearances and phenomenology are constantly changing as a result of the influence of perceptual relativity. An advocate of the third strategy must somehow account for this flux. The advocate hopes to do so by saying that subjects are visually aware of objective properties in *ways* that are constantly changing. An advocate also hopes to model these *ways* of being aware of objective properties by saying that the relation of visual awareness is a three-term relation, and that one of the terms is a package of the factors that are responsible for relativity.

As far as I can see, there is no hope of developing this view into a coherent package with sufficient explanatory power. According to the view, the objects of *visual* awareness are objective, intrinsic properties, but it is agreed on all sides that phenomenology is in constant flux. Thus, the view implies that we are not *visually* aware of phenomenology. How then do we know about phenomenology? If we aren't visually aware of it, we must be aware of it in some other way. An advocate of the view must, I think, maintain that awareness of phenomenology is a kind of second-order awareness of facts involving the three-termed relation of *visual* awareness. But what type of awareness could this second-order awareness be? It can't be visual awareness. Visual awareness is first order on the hypothesis we're considering. Nor can it be introspective awareness. Facts involving the three-term relation of visual awareness also constitutively involve factors like distance, angle of view, and lighting. These are all mind-independent properties and relations. They aren't accessible to introspection. It turns out, then, that the view is compelled to posit a form of awareness that is neither perceptual nor introspective. It is highly mysterious, relative to our current theories of awareness, and it should therefore be eschewed unless all other ways of approaching relativity fail.

The reader will notice a parallel between what I have just said about the third externalist strategy and my assessment of adverbialist accounts of awareness in later chapters. The strategy and adverbialism are closely related: they are both versions of the idea that perceptual relativity should not be explained in terms of objects of awareness, but rather in terms of modes of awareness or ways of being aware.

I turn now to internalist explanations of perceptual relativity.

The most famous of these is the *sense datum theory*; according to this view, perceptual experience constitutively involves awareness of mind-dependent objects and properties. But there have been other attempts to give internalist explanations. Instead of maintaining that the objects of perceptual experience are mental, *projectivists* have allowed that they are external, but have maintained that the properties that are attributed to them by experience are actually properties of mental states. On this view, perceptual experience results from a tendency to commit a certain sort of binding error, whereby representations of internal properties are erroneously bound to representations of external objects. I will argue that both of these views are mistaken.

According to standard versions of the sense datum theory, we are perceptually aware of objects and properties of two kinds. We are directly aware of sense data and their properties, and we are indirectly aware of external objects and their properties. Phenomenology is determined by objects and properties in the first domain, but the second domain is the one that we must navigate in order to achieve our goals. We grasp facts involving the second domain by drawing inferences from representations of the first domain.

The sense datum theory is importantly similar to the hypothesis that the objects of experiential awareness are retinal images, which we briefly considered in Section VI of the preceding chapter. That view was found wanting for two reasons. First, it cannot accommodate intuitions of perceptual transparency. We have the vivid impression that the objects of visual awareness are external to our bodies, and in fact stand at various distances from our bodies in three-dimensional space. Experiencing objects perceptually isn't at all like experiencing internal phenomena such as visual images or pains in our eyes. It would be highly counterintuitive to maintain otherwise. But to say just this is to underdescribe the situation. We have these intuitions because of facts of experience. Our *visual experience* seems to present objects of perceptual awareness as external to our bodies, and, in general, it seems to present us with objects arranged in a three-dimensional space. To say that the objects of visual experience are retinal images is to deny these experiential facts. It is therefore very hard to see how the retinal image theory could explain them. On the other hand, the view that visual representations have externalist contents can explain them nicely. Second, it is extremely plausible that the representational contents of perceptual experiences are determined by our biological need for information about the world—information that can guide navigation and support planning. Representations of retinal images could not serve this need. Hence, if experiential representations were concerned only with retinal images, they would have no immediate relevance to the informational needs that the visual system is designed to serve. I find this conclusion incredible.

Reflection shows that these objections to the retinal image hypothesis also count against the sense datum hypothesis, and that they have as much force with regard to the latter as they do with regard to the former.

I add that the motivation for the sense datum theory is extremely weak. It consists principally of the *argument from hallucination*, which runs as follows:

First premise: When we are subject to hallucinations it seems to us that we are perceiving a physical object with certain properties even though no such object is even remotely visible.

Second premise: It must be true in a hallucinatory situation that we are perceptually aware of *some* object with the relevant properties. It would be absurd to say that we are not aware of anything.

Lemma: Hence, it must be true in hallucinations that we are perceptually aware of objects that aren't physical. Mental objects are the obvious candidates. Let's call them *sense data*.

Third premise: Since hallucinations are indistinguishable from veridical perceptual states, it must be true that we are aware of sense data in veridical perception as well as its hallucinatory counterpart.

Conclusion: We are aware of sense data in perception.

From the perspective of representationalism, the force of this argument is negligible. The problem lies with the second premise. According to that premise, the only way to account for the fact that in hallucinations we seem to be aware of objects with certain properties is to suppose that hallucinations involve awareness of mental objects. But it is clear that there is an alternative explanation—namely, that when we hallucinate, we are simply *representing* physical objects with certain properties as existing. As we all know, it holds as a general rule that it is possible to represent objects that do not exist. Thus, for example, we all acknowledge that Bruegel's *Flight of Icarus* represents Icarus, a mythical character. Now it follows from these observations that anyone advancing the argument from hallucination would have to show that representationalist accounts of hallucinatory experience are wrong. But as far as I know, few people who have advanced a version of the argument for hallucination have even attempted to do this, and none who have attempted to do it have made a good showing. (There is a detailed representationalist account of hallucinations in the next chapter.)

I turn now to projectivism, which tries to explain relativity by maintaining that visual experience is the product of a systematic tendency to commit binding errors. More specifically, it maintains that every experience is formed by binding representations of properties of mental states to representations of external objects. My main complaint about this view is that it suffers from a kind of

inconsistency. There is good reason to think that the visual processing of information about properties makes use of information of the same kinds as the visual processing of information about objects. In particular, there is reason to think that both kinds of processing make use of information about depth. Because of this, there is a kind of inconsistency in maintaining that visual awareness of properties is concerned with properties of internal states while visual awareness of objects is concerned with fully external entities. I will illustrate this doubt about projectivism with two examples.

As we noticed in Section V of the preceding chapter, even though visual experience of size falls short of full constancy, there is generally an approximation to constancy, and this is due to computations that make use of information about the egocentric distances of objects. But information about egocentric distance is what enables the visual system to "externalize" objects by assigning them locations in physical space. So the processes responsible for the representation of size are bound up with the processes responsible for externalization.

Recent work on the experience of the lightness (that is, the apparent reflectance) of surfaces affords another example of the influence of information about 3D layout on our experience of properties. This work, most of it done or inspired by Alan Gilchrist, indicates that the lightness of an item depends on the depth of the surface with which the item is associated.

In a brilliant experiment conducted in the 1970s, Gilchrist manipulated the apparent depth of a piece of gray paper, which I will refer to as "G." Throughout the experiment, the degree to which G was illuminated remained the same. However, the manipulations changed the viewing conditions in such a way that G first appeared to be coplanar with a certain surface S1 and then appeared to be coplanar with a different surface S2. The surface S1 had a much higher reflectance than the surface S2, and it was also more brightly illuminated. Thus, G was first seen as coplanar with a surface that had a much greater luminance than it did, and then seen as coplanar with a surface which had a much lower luminance. (Luminance is the intensity of light that a surface reflects.) Gilchrist found that the lightness of G, its perceived or apparent reflectance, was hugely affected by the luminance of the surface with which it was seen to be coplanar. G seemed almost black when it appeared to be coplanar with S1, and almost white when it was perceived to be coplanar with S2. Thus, the lightness of G was heavily influenced by its apparent depth (Gilchrist 1977).

These studies led Gilchrist to propose the following *Coplanarity Principle*: In general, the lightness of an item depends on the ratio between the luminance of that item and that of the surface with which it is perceived to be coplanar. There is now considerable support for this principle. Gilchrist's experiments have been replicated many times, and evidence of a number of other kinds has been adduced. (For a review, see Gilchrist 2006.)

In view of these considerations involving the awareness of size and the awareness of lightness, we can conclude that the visual system takes account of the

same sorts of information about the environment in producing representations of properties as in producing representations of objects. It follows that there is a tension involved in maintaining, as projectivism does, that perceptual awareness of objects puts us in touch with external reality, but that perceptual awareness of properties is concerned with an internal domain.

III. Aspects of Perceptual Relativity Due to Peripheral Sense Organs

In Chapter 2 we noticed that there are three sources of perceptual relativity—(i) external factors like distance that influence proximal stimuli, (ii) features of the peripheral sense organs that respond to proximal stimuli, and (iii) internal factors like attention and adaptation. In this section I will briefly consider the influence of peripheral sense organs.

One type of receptor-caused relativity has to do with the spacing of photosensitive cells in the retina. Photosensitive cells are densely packed in the fovea, but their density decreases dramatically as they become increasingly peripheral. As a result, highly peripheral cells are significantly less capable of registering precise information about such features of objects as size, shape, location, and internal mereological structure. Now these differences in retinal capacities correspond to significant experiential differences. As introspection amply attests, the apparent sizes, shapes, and locations of objects in the center of the visual field are much more precise than the corresponding features of objects in the more remote regions of the visual field. Moreover, the mereological structures of objects are experienced in much greater detail. (For an illuminating discussion of retinal sensitivities at different eccentricities, see Block (2015).)

This type of perceptual relativity is an intrasubjective affair: equivalent stimuli presented to a single individual at different retinal locations are processed differently by retinal cells and therefore give rise to different appearances. Features of sense organs also give rise to intersubjective forms of relativity. For example, there are sex-linked differences in the wavelength sensitivities in cone cells. The peak sensitivities of middle wavelength cone cells in women tend to be different than the peak sensitivities of the corresponding cells in men, as do the peak sensitivities of long wavelength cells. As a result of these differences, the color experiences of women and men tend to be different, a fact that is manifested by differences in how subjects respond when asked to match colors (Neitz and Jacobs 1986, Neitz et al. 1993, Murray et al. 2012).

Here, then, are two additional ways in which appearances can fail to correspond to reality. What do they tell us about the nature of appearances? The theory of Thouless properties developed earlier does not directly address this question, for it is principally concerned with aspects of relativity that are due to partial constancy transformations. Those aspects are quite different than the ones that we are

considering here. But the general background picture I've been assuming can easily accommodate the present aspects. According to this background picture, it is the job of the visual system to compute functions of functions of . . . functions of proximal stimuli, where proximal stimuli are taken to be viewpoint-dependent properties of external objects. Now it is part of this picture that the initial computations are made by retinal cells. The inputs to these cells are quantitative properties of various kinds, and the outputs are also quantitative properties, so it is natural to think of the cells as computing functions. Moreover, the differences in retinal cells that we have been considering mean that the computed functions are different. Banks of retinal cells at different distances from the fovea will perform different computations on proximal stimuli that are identical, as will cone cells with different wavelength sensitivities.

I am thinking of the quantitative proximal stimuli that are the inputs to these computations as viewpoint-dependent properties of external objects, not as packets of light that are poised before the eyes, or as retinal irradiations. This is the natural view of them, given that we are concerned in this inquiry with ecological relevance. The outputs of the computations will also be viewpoint-dependent properties, for the computations will *include* partial constancy transformations that will lend the outputs a Thouless dimension.

To summarize, the aspects of relativity that are due to peripheral sense organs are easily accommodated by the hypothesis that visual processing consists of series of computations that have viewpoint-dependent properties as initial inputs and viewpoint-dependent properties as experience-level outputs.

IV. Aspects of Perceptual Relativity Due to Attention

The doctrine of perceptual relativity includes the claim that perceptual appearances are shaped by internal factors. I will focus here on the influence of attention on appearances, but there are other internal factors that exert an influence. One of these is adaptation, which occurs when cells that have been responding vigorously to a stimulus over a comparatively long interval become desensitized to it. Another example involves the experience of color, which is influenced by individual differences in cortical processing of information about wavelengths. What I will say about attention will not be directly relevant to all of the internal influences on experience, but it will have a direct bearing on adaptation, which is similar to attention in that it effects a change in the representation that is being used to characterize an object.

In addition to the introspective evidence that was available to earlier philosophers, we now have strong experimental evidence concerning the ways in which visual experience is shaped by attention. The work of Marissa Carrasco is particularly relevant here (Carrasco et al. 2004, Carrasco 2011). In one widely known

experiment, a subject *S* is confronted with two Gabor patches with different degrees of contrast, one on the left and one on the right. (A Gabor patch is a grating consisting of alternating dark and light lines. The degree of contrast of a patch is the average difference in luminance between the darker lines and the lighter lines.) Let P1 be the patch on *S*'s left, let P2 be the patch on *S*'s right, and let these patches have contrasts of 22% and 28%, respectively. The patches look appropriately different to *S* when they are viewed non-attentively. Carrasco showed, however, that when *S*'s attention is drawn to P1 by a brief preliminary stimulus, P1 and P2 appear to *S* to be equal in contrast.

Elaborating, Carrasco's result can be characterized as follows: At time T1, when *S* is viewing both P1 and P2 non-attentively, the two patches seem to have different levels of contrast; but at time T2, when S is attending to P1 but not to P2, P1 and P2 look to S to be the same in point of contrast. By the same token, P1 looks different to S when seen at T1 than it does when it is seen at T2. That is, *S*'s experience of P1 undergoes a phenomenological change between T1 and T2—a change that is due to attention.

This can be put by saying that *attention boosts apparent contrast*. Carrasco has also shown that attention boosts apparent size, apparent velocity, and apparent saturation of colors.

Here is a passage indicating how Carrasco interprets her results concerning apparent contrast:

> The visual system operates on the retinal image so as to maximize its usefulness to the perceiver, often producing nonveridical percepts. The visual system does not provide an internal one-to-one copy of the external visual world; rather, it optimizes processing resources. Attention is a pervasive example of this perceptual optimization: attention augments perception by optimizing our representation of sensory input and by emphasizing relevant details at the expense of a faithful representation. . . . (Carrasco 2011, 1496)

Carrasco evidently sees attention as introducing discrepancies between appearance and objective, physical reality, and she clearly has a Panglossian take on those discrepancies. As in other cases involving discrepancies, the discrepancies introduced by attention can be useful to the organism, provided that they can be calibrated to reality by the cognitive and motor systems. (For example, consider flint chipping. If attention makes facets of a piece of flint appear slightly larger than they are, that will enhance the chipper's ability to determine where the next stroke should fall.)

Can the influence of attention on appearances be accommodated within the present framework? Yes. Consider what happens when light that carries information about a certain level of contrast arrives at the eye. It will trigger a computational process that will eventuate in a representation. This is not the whole story, however, for the process will be different depending on whether or not it is shaped

by attention. If attention is not involved, the computations may implement functions f1, f2, . . . , fj and eventuate in a representation R, but if attention is involved, the relevant functions may be g1, g2, . . . , gi and the process may result in a representation R* with a different content. This presumably is what happens when attention changes the appearance of the left patch P1 in Carrasco's experiment. The same input results in different outputs because different processes are involved. So far so good, but what can be said about the fact that attention causes the left patch P1 to appear equal in point of contrast to the right patch P2? The details are not known, but it is clear that attentive processing of the input from P1 will involve different computations than non-attentive processing of the input from P2. Also, it is a familiar fact that different functions applied to different inputs can yield the same output. Thus, for example, 2^x and 4^y have the same outputs when x and y are set equal, respectively, to 4 and 2. Applying this moral to the case at hand, we can see that attentive processing of a viewpoint-dependent property involving P1 could issue in the same representation as non-attentive processing of a viewpoint-dependent property involving P2. The representation would be the same and the viewpoint-dependent property that serves as its content would be the same, so this picture can explain the fact that the appearances are the same.

There has been some tendency to interpret Carrasco's results as showing that attention introduces a puzzling amount of representational error. This interpretation is forced upon us if we think that the representations that are the outputs of attentive and non-attentive processing stand for objective, physical contrasts as opposed to viewpoint-dependent properties, for in Carrasco's example the only objective contrasts available to serve as contents of representations are 22% and 28%. Suppose that the representation R that results from attentive processing of information from P1 has 28% as its content, and suppose also that this same representation R is produced by the non-attentive processing of information about P2. Then we can explain why P1 and P2 look the same to subjects, but we're forced to say that R misrepresents P1. It is built into the example that the objective contrast of P1 is 22%. If, however, we think of R as having a viewpoint-dependent property as its content, we need not suppose that its content must be identical with either of the objective contrasts we have been considering. And, by the same token, we can avoid the conclusion that its deployment creates an error. Similar points apply to other types of attentional enhancement, such as enhancement of size.

V. Thouless Properties and Causal Theories of Representation

According to the picture sketched in the previous chapter, the properties that are represented in visual experience are Thouless properties—that is, viewpoint-dependent properties that are functions of more basic viewpoint-dependent

properties such as visual angles and angles of declination. We can represent Thouless properties experientially because the visual system is equipped with mechanisms that can compute the relevant functions. Now, when we reflect, we see that there is little room in this picture for the view that the properties represented in experience are among the causes of experiences. The causal factors specified by the picture are basic viewpoint-dependent properties like visual angles and angles of declination. (An angle of declination is a viewpoint-dependent property of an object X that is determined by the relation between the eye of a viewer and the point on the ground where X is resting. See Figure 2.4 in the preceding chapter.) It is properties of this sort that leave their imprints on the retina and thereby initiate the series of computations that lead to representations of Thouless properties. Thouless properties are viewpoint-dependent properties of external objects, just as visual angles and angles of declination are, but they do not themselves independently trigger retinal activity. They figure in perceptual processing only because computational mechanisms eventually generate representations that covary with them. It is co-variation rather than causation that links Thouless properties with the internal states that, according to the Thouless picture, represent them. (For an example of co-variation without causation, consider the relation that obtains between A, B, and C when A is a common cause of B and C. In a situation of this kind, C may covary with B but B is not a cause of C.)

Is this a problem? It is if we adopt the view, which appears to be widely shared, that properties that figure in the representational content of a perceptual representation must be causes of the representation. There are two main reasons for holding this view. I will describe these reasons and argue that the considerations on which they are based can be accommodated without supposing that there is an internal relation between causation and content.

One reason begins with the fact that representation seems to involve information. Although there are of course a number of different theories of representation, most of them endorse the idea that in a broad sweep of cases, at least, a representation of X must encode information about X (Dretske 1995, Burge 2010, Neander 2017, Shea 2018). The second step of this line of thought is the claim that causal accounts of information are superior to others. To see the motivation for this step, consider Neander's causal account of information in relation to the idea that information consists in probability-raising or reduction of uncertainty. According to Neander, "one event carries information about another if the second causes the first." Amplifying this remark, she says that an R-type event carries the information that a C-type event has occurred only if the C-event caused the R-event, *and* the C-event did so in virtue of its C-ness. This analysis has the great merit of honoring the intuition that information is *factive*: intuitively speaking, it is impossible for an R-event to carry the information that a C-event has occurred unless a C-event actually has occurred. Since a C-event must exist in order to cause an R-event, Neander's account of information can claim superiority to any

account which claims only that an R-event carries information about a C-event if the occurrence of the R-event merely raises the probability that a C-event has occurred.

In sum, there is an argument for an internal relation between content and causation that turns on two considerations: (i) the claim that representation involves information, and (ii) the claim that information must involve causation because information is factive.

There are, however, two problems with this argument. One is that it is possible to accommodate the factiveness of information without embracing a causal account. To accomplish this, it suffices to add a condition to so-called probability-raising accounts.

According to probability-raising accounts of information, whether an event of type T1 carries information about an event of type T2 has to do with whether the conditional probability **P**(T2/T1) is greater than the simple probability **P**(T2). Since Dretske 1981, accounts of information that are based on probability have played an important role in discussions of representation. (See, e.g., Shea 2018.)

Now, for the most part, theories of information of this sort don't entail factivity: events of Type T1 can raise the probability of events of Type T2 even if, in a given case, an event of the first type occurs without being accompanied by events of the second type. But there is a simple addition to such theories that brings them into line with the factivity requirement. Thus, instead of saying only that an R-event encodes information about a C-event if **P**(C/R) > **P**(C), a probability-raising theory can say instead that an R-event encodes information about a C-event if (a) **P**(C/R) > **P**(C) and (b) a C-event has actually occurred.

The other problem with the foregoing argument is that the factivity condition on information is a commonsense requirement that science seems to have left behind. When, for example, a biologist says that genes in zygotes carry information about the mature characteristics of an organism, it isn't built into the meaning of the claim that the organism will definitely exhibit those characteristics. The claim is meant to allow for the possibility of damage to the genes at later stages of development as a result of disease, and for impingement on development by cultural factors, such as the ancient Chinese practice of binding of women's feet to reduce their size. Evidently, biologists who attribute informational properties to genes do not have factive conceptions of information in mind.

These points about the impact of disease and culture on the efficacy of genes are counterexamples to purely causal theories of information. Reflection shows that there are also counterexamples of other kinds. Another type of counterexample involves the three-place relation of common causation. Intuitively speaking, a glimpse of a small grey animal climbing a tree carries the information that the animal has a bushy tail. This is because the property *being a squirrel* causes animals to be small, grey, and to have a penchant for climbing trees—and also causes

animals to have bushy tails. Moreover, the relation of information can link a property P1 and a property P2 even if there is no causal relation of any kind between P1 and P2. This would happen, for example, if two clocks were made by different artisans but were otherwise identical and the artisans happened to set the clocks to the same time at exactly the same moment. Laws of nature might guarantee that movements of the hands of these two clocks were perfectly correlated for many years; but still, it would be true that neither clock caused the behavior of the other and that neither their existence nor the synchronicity of their movements had a common cause. It is also common to speak of one event as carrying information about a future event even though the first event is a cause of the second event rather than the other way around. This is true in genetics, as we have seen, but it is also true in physics: physicists sometimes speak of a state of the universe as carrying information about later states.

These counterexamples to exclusively causal accounts of information suggest that we would be better off with pluralistic accounts, in which causal conceptions of information are combined with probabilistic conceptions.

I turn now to a second reason for thinking that there is an internal relation between content and causation. If the occurrence of a representation R raises the probability that one event has occurred, it also raises the probabilities of other events. Thus, if an occurrence of a certain representation R in the visual system raises the probability that an object in the environment is red, it also raises the probability that long wavelength light has been reflected by the object, the probability that cells in the retina that are sensitive to long wavelength light have been triggered, the probability that certain cells in the thalamus have been activated, and so on. But at the level of experience, at least, the contents of visual representations are unique. Experiential representations don't stand for activity in retinal cells or activity in cells located in the thalamus. Nor do they stand for the disjunction of the two. In view of this, it is clear that content is not determined by probability-raising. On the other hand, it seems that we have a chance of explaining content if we allow ourselves to speak of causation. Thus, we might say something like this: a property P is the content of a (visual) representation R only if, in normal circumstances, P imposes a structural characteristic C on the light L reflected by an object that possesses P such that L's having C is the dominant external cause of R's being tokened. This confers a privileged status on P that isn't shared by any property of events intervening between the object that possesses P and the corresponding retinal activity, and isn't shared by any events involved in the internal processing of the signal coming from the object that possesses P. It is possible to motivate the condition by pointing out that the properties that impose structural characteristics on light to which the visual systems are sensitive are usually possessed by objects that are germane to the needs and actions of humans and animals. This is true, for example, of colors, sizes, shapes, and textures. (I am here elaborating on a point made in Section VI of Chapter 2.)

At first sight, this line of thought appears to provide a good reason for thinking that the contents of representations must be properties that play a certain causal role in producing the representations. When we take a closer look, however, we appreciate that the line of thought only supports a much weaker conclusion—namely, that the properties that figure as contents of representations must *either* be identical with properties that play the relevant casual role *or* be coinstantiated with those properties (i.e., be instantiated by the same object). The second disjunct of this condition would be as successful in guaranteeing determinacy of content as the first disjunct, in the sense of singling out the same link in the relevant causal chain, and would therefore be as well protected against objections.

Now Thouless properties are in fact coinstantiated with the objective properties of distal objects that serve as triggers for the processes in the visual system that eventuate in experiential representations. That is, Thouless properties satisfy the second disjunct of the foregoing condition. This is one of their defining characteristics. Also, as the reader may recall, there is a good reason for supposing that it is Thouless properties, rather than properties of some other kind, that experiential representations signify. The reason is that Thouless properties are properties of precisely the objects that the visual system must represent if it is to perform its function of guiding behavior that fulfills needs. (See Chapter 1, Section V, and Chapter 2, Section VI.)

I conclude that both the factivity argument and the content-determinacy argument for causal theories of representation can be answered.

VI. Generalizing from Vision to Other Perceptual Modalities

We have thus far been considering a picture of vision and visual experience that consists principally of the following six claims:

1. There is a substantial divergence between the ways external objects appear to us visually and the ways they are in themselves.

2. This divergence between appearance and reality is not confined to special contexts that promote illusions, or to agents who suffer from visual deficiencies or disorders. On the contrary it is pervasive, characterizing even the visual experiences that occur in the most normal contexts. Moreover, the divergence takes a number of different forms and has a number of different causes, including differences of various kinds between the fovea and peripheral parts of the retina, and purely internal factors like attention and adaptation.

3. One of the principal causes of this divergence is the limitations of constancy transformations. To achieve full constancy, a transformation would have to compensate entirely for the variability and perspectivality of proximal stimulation, bringing the features represented by experience into perfect or near-perfect

alignment with their objective correlates. But the constancy transformations that the visual system actually employs fall far short of this standard. There are underconstancies of size, shape, color, location, and motion, and there are also overconstancies, as in the perception of slant.

4. To explain the ways that experience departs from constancy, it is necessary to suppose that the properties represented by experience are neither simple viewpoint-dependent properties like visual angles nor objective physical properties, but rather properties that R. H. Thouless called a "compromise" between properties of these two kinds. These properties are obtained by applying functions to quantitative features of proximal stimuli: more specifically, they are properties of the form *being such as to present a proximal stimulus S with certain quantitative characteristics Q1, Q2, . . . , Qn such that F(Q1, . . . , Qn) = N.* The functions are ones that are actually computed by the visual system in the course of processing the information carried by proximal stimuli.

5. Although objective properties are not represented by visual experience, there are post-experiential representations of various kinds that do a better job of keeping track of them. These post-experiential representations are found in the cognitive system, the higher levels of the visual system, and the motor system.

6. It can be tempting to assume that the discrepancies between visual appearances and the corresponding external facts are due to failings of the visual system, imperfections that evolution has so far been unable to correct. In fact, however, at least in some cases, the discrepancies admit of a rosier gloss. For example, as we saw in Chapter 2, when representations of Thouless sizes are combined with cues of various kinds, they permit us to recover information about objective distance and also information about objective sizes. Accordingly, systems of Thouless representations may be regarded as efficient systems for combining different streams of information, thereby making them simultaneously available to higher level processes that need to be apprised of multiple streams at a single time and a single place.

I will turn now to consider the extent to which these claims fit non-visual perceptual modalities. To anticipate, what we will find is that the fit is pretty close in the case of audition and the case of touch, but that in other cases, such as the contact perceptual systems of olfaction and taste, only some of the six claims apply. More specifically, we will find sharp divergences between appearance and reality in all of the cases we consider, but we won't always encounter counterparts of Thouless properties. There are no strong reasons for attributing constancy transformations to the contact systems.

In this inquiry I will focus on audition, touch, and olfaction. There will be no explicit discussion of taste, thermoperception, or proprioception, but much of what I will say about olfaction applies to taste, and there will be allusions to proprioception in the discussion of touch. I will not be able to consider any of the many fascinating perceptual modalities that animals possess but that Nature chose not to bestow on human beings.

VII. Appearance and Reality in Audition

To approach this issue, it is necessary to adopt a theory of the ontology of sounds. The theory I will presuppose is the one recommended in Casey O'Callaghan's excellent book on the topic (O'Callaghan 2007). One of the features of O'Callaghan's view is a sharp distinction between sounds and sound waves—sounds are located in vibrating objects, such as vocal cords, bells, and car horns, and soundwaves are compression waves in the atmosphere carrying information about such vibrations. There is a range of other theories that draw this distinction (e.g., Casati and Dokic 1994, Pasnau 1999, Kulvicki 2008), and it would be possible to develop the account I wish to propose in terms of any one of them. But for reasons that it is unnecessary to dwell on here, I find O'Callaghan's view the most plausible of these options. Casati et al. 2020 contains a balanced comparative assessment.

O'Callaghan sees sounds as events in which a vibrating object interacts with a medium and thereby transmits vibrations to it. For example, the sound of a tuning fork is the event of the tuning fork's disturbing the surrounding air. In other cases, sounds involve the interaction of two or more objects—perhaps a collision of two cars or a strumming of a guitar's strings. But in all cases, "particular sounds are events of oscillating or interacting bodies disturbing or setting a surrounding medium into wave motion" (p. 60). Thus, sounds are relational events—in a word, *disturbances*—that are located at the points in space where the events occur.

On this account, the properties we associate with sounds, such as loudness, are grounded in physical properties of disturbances of the given sort. Loudness itself is grounded in the intensities of disturbances, and pitch is grounded in the frequencies of disturbances. Timbre is more complex, but it too has a physical ground, involving multiple frequencies together with amplitudes at the onsets and offsets of sounds. (Timbre is usually explained in terms of an example, by saying that it is what distinguishes the sounds produced by different musical instruments that are playing the same note at the same volume.)

When understood as disturbances, sounds contrast sharply with sound waves, which O'Callaghan identifies with progressions of vibratory behavior through media, carrying information about properties of their sources like volume and pitch. A soundwave is something that happens to a medium (p. 26), but it is in turn affected by features of the medium, especially by the medium's density and elasticity (p. 25).

There is a tradition of identifying sounds with waves rather than the vibratory causes of waves. O'Callaghan gives strong reasons for preferring his different view. One is that if sounds were soundwaves, it would be true that our experience of the locations of sounds would be systematically deceptive. I experience the sound of the tuba as located where the tuba is located, not as having multiple locations at different times (as would be the case if sounds were waves), or as

spread around in space. Another reason for doubting the wave theory of sounds is that it conflicts with our experience of the durations of sounds. The jarring noise of a car crash lasts but a moment, but the wave emanating from the crash continues on for an extended period. (These objections to the wave theory of sounds count in favor of all theories that distinguish between sounds and soundwaves, not just O'Callaghan's account.)

I endorse all of the foregoing components of O'Callaghan's position and will assume them in what follows. I will also assume his view that that it is sounds themselves, and not soundwaves, that are the direct objects of auditory perception (p. 60, p. 69).

So much for background views. We should now observe that there is a broad range of discrepancies between the ways sounds appear to us and the ways sounds are in themselves. To begin with, as O'Callaghan points out, our perceptions of pitch are systematically distorted by Doppler effects (pp. 103–9). (Doppler effects are shifts in the pitch of sounds due to the relative motion of sound-producing objects and observers. A classic example is the shift to a higher pitch in an ambulance siren as it approaches an observer. Due to the forward motion of the ambulance, the distances between soundwaves are compressed, which entails an increase in frequency. Since frequency determines pitch, an increase in pitch accompanies this compression. Doppler effects can be due either to the motion of the object generating the sound (source motion) or to the motion of the observer (subject motion).) Discussing Doppler effects, O'Callaghan writes as follows:

> Both source motion and subject motion produce illusions of altered pitch thanks to how waves transmit information about sounds and excite auditory experiences. In neither case does a sound alter its qualities due to relative motion of source and subject. Rather, a sound merely seems to have altered its pitch thanks to such relative motion. (p. 108)

Now O'Callaghan thinks that events involving Doppler effects are a special case: in most cases, he maintains, our perception of the qualities of sounds is veridical (p. 164). But I think we must recognize that there are a number of other discrepancies between the objective properties of sounds and the ways sounds appear to us. For one thing, while pitch is a monotone increasing function of frequency, it isn't a linear function (Wolfe et al. 2006, 228–9; for a graph of the function, see Coren and Ward 1984, 202). But also, there are discrepancies between the perceived qualities of sounds and their objective properties in realms of auditory perception other than pitch. To illustrate, it is a truism that the perceived volume of a sound decreases with distance: a car horn that seems jarringly loud to passengers in the car seems to be of moderate volume when perceived by pedestrians half a block away, and is barely noticeable when perceived from a distance of three blocks. Perceived volume is characterized by a steep underconstancy.

Similarly, and for related reasons, underconstancy is a feature of our experience of the distance of sounds (Coleman 1962, Loomis et al. 1998, Kolarik et al. 2016). Appearances are also misleading with respect to directionality (Middlebrooks and Green 1991).

How are these discrepancies to be explained? One option is to maintain that auditory perception is broadly and systematically illusory—that audition systematically attributes objective properties to objects that they do not possess. Views of this sort are perhaps the ones that first come to mind, but I think it is good policy to keep illusions to the absolute minimum. As we observed in Section I, a perceptual system that was systematically erroneous in its attributions of objective properties to objects would pose a grave threat to the agents who possessed it, because it would cause agents to form false beliefs about the external phenomena on which the success or failure of their actions depend. Back in Section I we were considering vision, but the point is also quite relevant to our present concerns. To see this, consider an agent who hears a soft and gentle thrashing in the bushes. Suppose that the agent's auditory system represents the thrashing sound as being objectively soft and gentle. This would presumably lead the agent to believe that the sound was produced by a comparatively small animal that poses no physical threat. But this belief could be quite misleading, for the sound could in fact have been produced by a large animal that was located at a distant waterhole. By the same token, the belief could be quite dangerous, leading the agent to ignore the sound in planning actions. It could, for example, lead her to head to the waterhole for a drink. The agent would have been much better equipped if her auditory system had represented the sound as having a more complex property than *thrashing that is objectively soft and gentle*—-a property that left room for interpretation by higher level cognitive systems, such as the property *thrashing that seems soft and gentle*.

Theories of auditory experience are therefore confronted with a problem. On the one hand, there is a great deal of phenomenological and experimental evidence indicating that there are substantial discrepancies between the ways sounds appear to us and the ways sounds are in themselves. But, on the other hand, we don't want to explain these discrepancies by saying that auditory experience systematically misattributes objective properties to sounds. How then can the discrepancies be explained? There is a range of options, but two are especially promising. (i) According to the first of these alternatives, the properties of sounds that we directly perceive are viewpoint-dependent properties that are counterparts of visual angles—properties like *causing a soundwave that has intensity I upon its arrival at spatial location L* and *causing a soundwave that has frequency F upon its arrival at L*. That is to say, it is properties of this sort that are represented experientially. This idea affords an easy explanation of the failures of constancy we noticed a moment ago. (ii) The second alternative is the idea that we directly perceive counterparts of Thouless properties—that is, properties that the auditory system comes to represent by computing functions of properties of type (i), but

that are subject to underconstancies of various kinds. For the reason given in the preceding paragraph, both (i) and (ii) seem more attractive than the idea that the auditory system systematically misconstrues objects by attributing objective properties to them that they do not really possess.

Although the first alternative can be tempting, the second alternative seems to provide the best overall fit with the data. This is because the departures from constancy are limited by constancy transformations of various kinds. To illustrate, Bronkorst and Houtgast (1999) found a tendency toward location constancy when subjects were exposed to sounds in enclosed spaces, and several investigators (Zahoric and Wightman (2001), Altman et al. (2013)) have obtained similar results concerning loudness constancy. (Subjects have access to an additional cue concerning location and loudness in enclosed spaces. This cue involves reverberant energy from sounds reflected by floors, walls, and ceilings. See Kolarik et al. 2016 for a review.) There are also constancy effects for pitch (Warrier and Zatorre 2004) and timbre (Sandell and Chronopoulos 1997), and they are rife in the perception of speech, particularly with respect to phonemic character. (See Goldstein 2010, 315–16, for a summary.) Collectively, these findings indicate that the auditory system has multiple mechanisms that perform constancy transformations on proximal properties, and produce representations of objects that reflect the results of these transformations.

These constancy transformations go some way toward compensating for the effects of such factors as distance and relative motion, but they do not achieve full constancy. Underconstancy remains: even in an enclosed space, the perceived volume of a sound produced a few feet away is greater than the perceived volume of an equivalent sound at a distance of 100 feet, and the perceived pitch of the whistle of an oncoming train is higher at a distance of 100 yards than it was a moment earlier, when the train was a quarter mile away. As is the case with vision, the constancy transformations for audition are *partial* transformations. This is a good thing, as is illustrated by the forgoing example involving thrashing noises.

To summarize, in both visual and auditory perception, there is a pervasive contrast between appearance and reality, due to the variability and perspectivality of proximal stimuli. Also in both cases, perceptual experience is shaped by partial constancy transformations—transformations that partially compensate for, but do not eliminate, the gaps between proximal stimuli and their distal correlates. The result for auditory perception is that the features of sounds that are represented by auditory experience are relational, viewpoint-dependent properties, obtained by applying computable functions to proximal stimuli. These properties are not simple properties like volume and pitch, but rather, beneficially, more complex properties that support inferences about the simple properties. This conclusion parallels our main conclusion about visual perception.

It is important to recognize, however, that constancies and failures of constancy are not nearly as well studied in auditory science as they are in vision science.

The picture that I have offered here is but a sketch. It is clearly desirable to have a more detailed and more nuanced account of these topics, but that must await further scientific inquiry.

VIII. Appearance and Reality in Touch

What about touch? Does tactual experience resemble visual experience in being characterized by profound differences between appearance and reality? And if so, to what extent can the apparatus of Chapter 2 be used to explain these differences? In particular, is it helpful to think of the tactual system as computing functions of functions of . . . of functions of viewpoint-dependent properties, and as arriving thereby at experiential representations of properties that to some degree abstract away from the variability and perspectivality of their viewpoint-dependent inputs? I will maintain that the apparatus of Chapter 2 does indeed help us to understand tactual experience, while acknowledging that in the case of touch there are fewer experimental studies pertinent to determining the degree of fit of that apparatus than there are in the case of vision.

I am going to assume that the objects of tactual awareness are external objects that in one way or another come in contact with the skin. I am also going to assume that there is a broad range of properties of such objects that can activate touch and that are ultimately responsible for tactual experience. These properties include size, shape, texture, weight, viscosity, pressure, hardness, solidity, vibration, temperature, motion, distance in egocentric space, direction in egocentric space, and length and breadth of the facing surfaces of objects laid out along the skin. It is widely assumed, in both the philosophical and psychological literatures, that it is the function of tactual experience to represent these properties, which are of course objective, physical characteristics. I will be challenging this view, maintaining instead that experience represents properties that are appearances of objective properties rather than objective properties themselves; but I believe that experiential representations of appearances combine with other mental resources, such as memory and cognition, to provide us with fairly robust access to objective sizes, objective shapes, objective textures, and so on.

There is a third view about the objects of tactual experience that contrasts sharply both with the standard view and with the alternative that I wish to defend. This is the idea that the objects of tactual experience are bodily sensations and properties of sensations. In other words, according to this third view, we should assimilate tactual experience to experience of bodily states such as pains, itches, and tingles. There is a lot to be said in favor of this view, especially in terms of tactual phenomenology. Touch your left hand with your right index finger. Doesn't your finger seem to be causing phenomenological changes inside your left hand? Equally, don't you experience a sensation in your left bicep if you use

your right hand to pinch it? Further, if a masseur were to rub or knead your back, wouldn't that give rise to a series of internal sensations, perhaps characterized by a release of tension, or by a glow of pleasure? It would be odd to describe experiences of these kinds as representing external objects. As Matthew Fulkerson says in his thought-provoking book (Fulkerson 2014, 8), "touch often involves a seemingly subjective experience of the present state of our own bodies *from the inside*."

It can be tempting to think of this third view as in competition with the standard view, and also with the alternative I will recommend, but I think that a better way to accommodate the phenomenological data is to suppose that there are two systems of representation involved in touch, one that is used to track bodily sensations and another that is used to track properties of external objects that are in contact with the body. On some occasions only the first of these systems is activated; these are usually occasions when one passively receives stimulation from outside. On others the second is active; these are almost always occasions when one is actively using touch to explore objects. But as Fulkerson points out (p. 9), there are also occasions when both are activated: "Some experiences, like dipping our fingers in warm water, seem to have both a body-directed and an object-directed character. We experience the warmth *of the water*, but we also experience *our finger being warmed*." A further point is that it seems possible in some cases to toggle back and forth between experiencing a bodily sensation and experiencing an external object by shifting our attention.

There is, then, a distinction between passive touch and active tactual exploration of external objects. This distinction has received a great deal of attention in the literature, albeit under different names. Lederman and Klatzky (2009, 1439) distinguish between cutaneous and kinaesthetic modes of tactual awareness; Gardner and Johnson (2013, 520) distinguish between passive and active modes of awareness; and Fulkerson (2014, 12) and Matthen (2021, 197) distinguish between a tactile mode and a haptic mode. Adopting Fulkerson's terminology, the difference can be explained by saying that while tactile awareness only requires information from various kinds of mechanoreceptors in the skin, haptic awareness rests both on information from mechanoreceptors and proprioceptive information about the locations of various parts of the body. Information indicative of top-down control is highly relevant too (Gardner 2021), as are kinesthetic information and visual information. Thus, for example, when you are grasping a donut and tactually exploring its surface, with a view to determining its shape, your awareness of its curvature crucially depends on proprioceptive and kinesthetic information concerning the curved trajectory of your hand. Indeed, it seems that one or more kinds of information beyond tactile sensation are required for all forms of tactual awareness that count as externally directed. (Like Fulkerson, I will use "tactual" as a general term that encompasses both tact*ile* and haptic awareness.)

There are various theories of how tactile information is integrated with other kinds of information to give rise to haptic perception. Here is one such account,

in the words of Esther Gardner, which emphasizes the relationship between tactile information and information indicative of top-down control:

> During active touch, descending fibers from motor centers of the cerebral cortex terminate on interneurons in the medial dorsal horn that receive tactile information from the skin. Similar fibers from cortical motor neurons terminate in the dorsal column nuclei, providing an *efference copy* (or *corollary discharge*) of the motor commands that generate behavior. In this manner, tactile signals from the hand resulting from active hand movements may be distinguished centrally from passively applied stimuli in the neurological exam or in psychophysical tests. (Gardner 2021, 436)

(Efference copies are explained in Appendix I of Chapter 5.) It seems unlikely that this is the whole story, since it makes no mention of proprioceptive, kinesthetic, and visual information, but it gives us a piece of the puzzle.

Although it is important to acknowledge the duality of tactual experience, I will be concerned here with touch qua faculty of external awareness. Chapter 6 is concerned with awareness of pain, and much that I say there can be applied to tactual awareness of other bodily sensations. (For further discussion of the duality, see Fulkerson pp. 100–4 and Matthen 2021.)

I turn now to the question of how well tactual experiences reflect external reality. I will begin by reviewing a few illusions that may be familiar to the reader—they are often cited in texts on perception. The examples I have chosen are not isolated effects, occurring only in special contexts. They illustrate general, systematic tendencies in tactual processing to produce appearances that diverge from reality. I will then move on to examples more closely tied to experimental work.

The *tau illusion* involves successive stimuli (e.g., taps) applied to parts of the body separated by an objective distance D. The perceived length of D depends on the temporal interval between the stimuli: D is perceived as being shorter when the second stimulus follows hard upon the first, and as being longer when there is a larger temporal gap between the stimuli (Helson 1930). This effect is complemented by the *kappa illusion*, in which the perceived length of a temporal interval T between two stimuli varies as a function of the spatial distance between the parts of the body to which the stimuli are applied (Cohen et al. 1953). The *horizontal–vertical illusion* parallels a familiar visual illusion. It occurs when two lines of equal length are arranged to form an inverted T (⊥). Haptic exploration of the lines leads to the erroneous impression that the vertical line is longer than the horizontal line (Gentaz and Hatwell 2004). As these authors point out, this illusion is probably related to the tau effect. Apparently it takes longer for subjects who are haptically exploring an inverted T to trace the vertical line than it does to trace the bottom, horizontal line, despite the fact that the lines are the same

length. In line with the tau illusion, this leads them to experience the vertical line as longer than its horizontal partner. The *size–weight illusion* is due to a distorting influence of vision on haptic perception of weight. This illusion occurs when a subject is shown two objects of somewhat different size and then asked to determine the weights of the objects by lifting them. The weights are identical, but it nonetheless seems to the subject that the smaller object weighs more (Buckingham et al. 2016). A fifth example is the familiar *temperature illusion*. In the classic version, this illusion involves three pails of water, one hot, another cold, and a third that is of an intermediate temperature. A subject places one hand in the hot water and the other hand in the cold water, and then, after a minute during which the subject's hands respectively adapt to these different temperatures, places both hands in the water of intermediate temperature. It will seem that the heated hand is now in cold water and that the cooled hand is now in hot water (Berkeley 1953).

Experimental work has uncovered a number of systematic discrepancies between tactual appearances and external reality. There has, for example, been considerable work on the perceived sizes of objects that are laid across a stretch of skin. As one would expect, given that mechanoreceptors near the surface of the body are more densely packed in some areas than in others, the perceived size of an object is highly dependent on the location of the stretch of skin where it is positioned, with, say, objects located on the index finger seeming significantly larger than objects located on the upper arm, even though the objects are of equal size (Lederman and Klatzky 2009). Somewhat surprisingly, however, this effect is damped by visual awareness of the two affected regions. That is, there is a constancy transformation that integrates tactile and visual information and thereby reduces the size of the effect, though it does not eliminate it. (It is another example of a *partial* constancy transformation.) (Taylor-Clarke et al. 2004)

There are also systematic divergences between appearance and reality in other areas of spatial awareness. One such divergence involves the perceived orientations of objects. In a series of experiments, Kappers and Koenderink (1999) asked subjects to move bars placed on a horizontal surface in such a way that their orientations satisfied certain criteria. In some experiments, the goal was to match the orientation of a reference bar, which had been established haptically, by moving another bar located at various distances from the first. In other experiments, the goal was to point a bar in the direction of a marker that had previously been touched. The investigators found that "In all experiments large systematic deviations (up to 40°) were made" (p. 781). Kappers later (2007) extended this finding to situations in which bars were located on vertical planes. Seeking an explanation for all of these results about orientation, Kappers conjectured that they occur because subjects are trying to perform tasks that are defined by allocentric reference frames (for example, by the surface of a table on which a task is being performed). They fail because their efforts are constrained by the egocentric reference

frames that are more fundamental for haptic exploration. That is, they wind up using a reference frame that is not entirely allocentric, but is rather an average of egocentric and allocentric frames (Kappers 2007, Lederman and Klatzky 2009).

Perceived distances and directions in 3-D space also diverge from their objective correlates, and the same is true of perceived 3-D shapes. There are many results in these areas. Here is Lederman and Klatsky's summary of this literature (2009, 1446):

> The perception of geometric properties beyond fingertip scale is subject to a number of influences that undermine veridicality in systematic ways. For example, curvature perception depends on whether the curvature is convex or concave . . . , the direction of movement over the surface . . . , the position of the stimulus on the hand . . . , and on shape features other than the judged curvature . . . Haptic perception of linear extent is affected by the path length, curvature . . . , rate of exploration between endpoints . . . , and other linear elements in the field.

It should be noted that there are a number of important results not mentioned in this passage. To illustrate, Klatzky and Lederman found (2003) that subjects have difficulty returning to a point in space that they have previously touched, and Hartcher-O'Brien et al. (2014) showed that perceived curvature varies with distance from the observer. Hartcher-O'Brien and colleagues also established that these distance effects are damped by two modulatory processes, but found that the processes in question result in underconstancies and overconstancies.

Moving away from perception of spatial properties, there is a sizeable literature on the tactual perception of texture, much of it focusing on roughness (e.g., Tiest and Kappers 2007, Lederman and Klatzky 2009, Yoshioka et al. 2011). One important finding (Yoshioka et al. 2011) is that there are strong constancy transformations associated with the haptic experience of roughness, but that they appear to play no role in the tactile experience of that magnitude.

To summarize, although touch is a contact sense, and there is therefore no contrast between distal and proximal stimuli, there is nevertheless a vast divergence between tactual appearances and the objective properties of tactual stimuli. This divergence has a number of different facets. Correlatively, it has a number of different causes, including characteristics of the mechanoreceptors that record stimuli, characteristics of our proprioceptive equipment, characteristics of internal processors, and interactions between such processors and faculties associated with other perceptual modalities, including especially vision. Constancy transformations play an important role in some aspects of tactile and haptic processing, and some of these transformations are known to be partial, their outputs departing significantly from perfect constancy. In view of these facts, it is natural to conclude that the properties represented by tactual experience are not objective

properties of external stimuli, such as objective sizes, shapes, weights, and textures, but rather relational properties of external stimuli—properties that are arrived at by applying computable functions to tactual and proprioceptive information concerning our interactions with such stimuli. It is this view that best explains the phenomenology of tactual experience, which definitely involves a strong impression of externality, but is also palpably out of alignment with objective reality. Could we suppose instead that objects with certain objective sizes, shapes, weights, textures, and so on are represented by tactual experience, but that tactual experience misrepresents their properties massively and systematically? We have considered hypotheses like this in connection with vision and audition and have seen that they have trouble explaining why human action tends to be successful. Here I will offer an objection of a different kind. Suppose that a representation R stands for an objective property P, but that it is rarely if ever used to attribute P to an object that actually possesses P. Instead it is always, or virtually always, used to attribute P to objects that in fact possess a different property P*. In virtue of what fact or facts does R stand for P? What keeps us from saying instead that it stands for P*? In general, how can a representation acquire a content if it bears no privileged causal or correlational relations to that content? Since there is no evident way of answering these questions, there is a substantial presumption in favor of taking representations to stand for the relational properties with which they are strongly correlated.

I concluded the last section by observing that comparatively little work has thus far been done on constancy transformations and departures from constancy in connection with audition. A similar remark applies here. Most of the story about tactual constancies remains to be told. Hopefully it will be possible to make more detailed comparisons of visual appearances with tactual appearances in the near future.

IX. Appearance and Reality in Olfaction

The goal in this section is to determine whether—and, if so, to what extent—the account of visual experience developed in Chapter 2 can be transposed to olfactory experience.

Before undertaking that inquiry, however, we should consider the question of whether it is appropriate to investigate olfaction in a book that is concerned with *perceptual* experience. Although olfaction is widely regarded as a form of perception, this view is by no means universally held. It is opposed, for example, by Christopher Peacocke, who seems to think of olfaction as purely sensory. ("A sensation of . . . [smell] may have no representational content of any sort, though of course the sensation will be of a distinctive kind" (Peacocke 1984).) Tyler Burge is another opponent:

> Under certain limited conditions, location constancy through single or relatively short-term sensory registrations, hence perceptual objectification, is in principle possible for olfaction. Such conditions may well be realized in certain species. I think, however, that most animals that locate food or other targets by smell do so by the following technique, which does not make use of perception. They first register direction through intensities on one or another side of the body. Then they use a homing or beaconing technique that serially follows directional intensity (determined by relative intensities on one or another side of the body) of a favored odor to the target. This technique is widespread. It enables numerous animals to use a non-perceptual sense to locate targets. (Burge 2010, 415)

Burge is moved to deny that olfaction is perceptual because it appears not to involve representations that assign properties to distal objects. Instead it simply registers compositions and concentrations of chemical mixtures that are in contact with one's body. Moreover, it appears not to involve constancy transformations—processes that abstract away from the mercurial nature of peripheral stimulation, so as to bring experiential representations in line with objective magnitudes. On the contrary, olfactory experience seems to be subject to fluctuations that parallel the fluctuations in peripheral sensory activity. Since Burge believes that perception constitutively involves representation of objective properties, and also that representation of such properties requires constancy transformations, this is for him a decisive reason to deny that olfaction is a genuinely perceptual faculty.

As I read him, then, Burge has two reasons for doubting that olfaction is perceptual. One is that it fails to assign distal positions to the objects of olfactory awareness, and the other is that olfaction does not involve constancy transformations, and therefore could not be concerned with objective reality. I don't find either of these reasons persuasive. First, it seems that olfaction does assign spatial positions to its objects: it represents them as present in one's nostrils, and this is no less a position in 3-D physical space than locations that lie at some distance. Perhaps Burge would say that externality is not enough: distality is required too. But why insist on distality? There is no such requirement in the case of touch. Second, there are two problems with Burge's constancy argument. One is that it appears that olfaction is in fact governed by certain limited constancy transformations (Millar 2019). More importantly, it seems quite wrong to suppose that objective perceptual reference requires constancy. It is enough that perceptual representations have veridicality conditions involving entities that are independent of the perceiving subject. And this requirement is fully satisfied in the case of the representations that figure in olfactory experience. Thus, olfactory experiences can be hallucinatory. Indeed, olfactory hallucinations are common: recent research by Kathleen Bainbridge and colleagues (Bainbridge et al. 2018) has shown that 6.5 per cent of Americans over 40 suffer to some degree from the

condition known as *phantosmia*, in which odors are hallucinated, sometimes quite vividly. The experiences of these people represent their nostrils as containing chemical compounds that are not actually present. There are also olfactory illusions. The reader may have experienced them personally, for the disorder known as *parosmia* is a frequent consequence of Covid-19. Here are three testimonials from victims of parosmia that I've found at a BBC News site (https://www.bbc.com/news/stories-55936729):

> (i) In spring we both caught Covid and he was hospitalised. I struggled down to the kitchen to make coffee and toast for myself. Exhausted by such a simple task, I clung to the ritual and pictured him beside me. We both recovered, but coffee and toast is now repulsive to me—like a field just sprayed with manure . . . unpleasant with a sweet fermented smell on top.
>
> (ii) I hesitated before I put my nose in the glass. I'd had the wine before, I knew how it should be. The bouquet was wonderful—honeyed, butter with peach and a hint of citrus. But then I took a sip and it hit me. The smell and taste of rotten, putrefying fruit came rushing in on the aftertaste. I felt sick.
>
> (iii) People ask me what smells or tastes I miss, but answering that is very difficult. After six months of living with parosmia, I don't miss any because I have forgotten what normal tastes and smells are like. Maybe it is my body's way of coping with what I've lost. The new smells seem to have imprinted on my brain permanently—a strong sharp chemical smell mixed with a potent rancid sewer smell that instantly makes my stomach turn.

In sum, there is extensive evidence that olfactory experiences can be both hallucinatory and illusory. There is no other natural way of describing the testimony of those who suffer from phantosmia and parosmia. But this entails that olfactory experiences have veridicality conditions that involve external phenomena. That seems an excellent reason for supposing that such experiences have objective import, whether or not they result from constancy transformations. (Batty (2010) would second my contention that there are hallucinations of smell, but she would deny that there are olfactory illusions. Perhaps she would respond to the testimony of parosmia victims by saying that odorants like the aromas of coffee and wine just smell different than they did in pre-Covid days—there is no basis for thinking that the post-Covid smells are inaccurate. As against this, I would urge that that the olfactory capacities of victims of parosmia have been *harmed* by Covid. The olfactory systems of the victims are no longer functioning properly. This seems a strong reason for thinking that they are serving up illusions.)

In addition to doubting that olfactory experiences have objective import, Burge has a similar doubt about taste (Burge 2010, 415–16). As I understand him, his reasons for this second form of skepticism strongly parallel his reasons for

skepticism about olfaction. If this is right, then the former reasons are subject to objections of the same sorts as the latter. Like smells, tastes have a location in 3-D space. They occur in the mouth. And again like smells, they are subject to hallucinations (Peng et al. 2015) and illusions (Spence et al. 2015). Thus, as Spence et al. point out, there is considerable evidence that the colors of drinks and foods can distort our experience of their tastes. As with olfaction, then, there is an appearance/reality distinction in the case of taste. (Speaking more generally, much of what I will say about olfaction in the present section applies with equal force to taste.)

It follows from what has been said so far that olfactory appearances can diverge sharply from reality. But the examples of divergence I have cited are quite special, in that they are due to disorders in our perceptual apparatus. They have no tendency to suggest a divergence of appearance from reality that parallels the systematic, pervasive discrepancies between appearance and reality that we found in the case of vision. Moreover, one might doubt that there is such a parallel. After all, in the case of vision, such factors as distance, angle of view, lighting, and relative motion make for very substantial differences between distal stimuli and proximal stimuli. There is a similar gap between distal stimuli and proximal stimuli in the case of hearing. In the case of olfaction, however, there is no comparable gap between the distal and proximal causes of experience. Olfaction is a *contact* perceptual modality, in the sense that the primary objects of olfactory awareness are present in the nostrils of the perceiver. In view of this, it can be tempting to suppose that the only discrepancies between appearance and reality in the case of olfaction are due to special factors like phantosmia and parosmia.

But this is just an initial impression. Further inquiry reveals a number of other factors. These include adaptation (an aspect of a complex odor may become invisible due to lengthy prior exposure to a simpler odor that presents only that aspect), attention, interactions with taste and other modalities, olfactory metamers (odorants composed of quite different molecules can give rise to the same olfactory experience), priming, simplification (odorants that are in fact highly complex can be experienced as simple), and the influence of learned associations and task demands. Further, as I will discuss at the end of this section, it turns out that there is a sense after all in which the objects of olfaction are governed by a proximal/distal distinction.

All of these factors are important. For the moment, however, I will focus on the ways olfactory appearances are shaped by learned associations, leaning heavily on the work of John McGann and his associates (Kass et al. 2013, Kass et al. 2016, McGann 2018).

One of the basic findings in this area concerns the sensory and neural changes that mice undergo when they are taught to expect an electric shock after being presented with a certain odorant. Inevitably, they begin to experience fear on smelling the odorant. This has the effect of enhancing their ability to discriminate

between that odorant and neighboring smells. More surprisingly, it enables mice to discriminate between odorants that they had previously been unable to distinguish. That is, there is behavioral evidence that fear conditioning changes the way that mice experience chemical compounds. This behavioral evidence is complemented by evidence of changes in the structure and physiology in the olfactory system, including changes in peripheral sensory neurons, changes in the olfactory bulb, and changes in the piriform cortex.

McGann's work on rodents complements research on human olfactory experience (Li et al. 2008, Áhs et al. 2013, Parma et al. 2015) and research on other perceptual modalities (Polley et al. 2004, Zhang et al. 2013, Knight et al. 1999, Song and Keil 2014). Pooling these findings with his own results, McGann summarizes the situation as follows:

> Associative learning can induce dramatic stimulus-specific plasticity in both the behavior of individual sensory neurons and the macroscopic remapping of sensory regions, including auditory, olfactory, somatosensory, gustatory, and visual systems. These effects are best understood in primary sensory cortices but also occur in precortical structures and even as early as the primary sensory neurons themselves . . . The function of these forms of plasticity may be to improve detection or discrimination of critical stimuli or to enable these stimuli to better engage attentional and learning circuitry. (McGann 218, 573–4)

McGann also notes that associative modulation of neural structures and olfactory experience is not confined to fear conditioning. Associative learning based on rewards achieves comparable effects.

It appears, then, that certain kinds of learning and conditioning can influence the way we experience olfactory stimuli. Experience of stimuli can change profoundly even though the stimuli themselves remain the same. We have also taken note of factors of several other kinds, such as adaptation and attention, that can influence olfactory experience, causing it to diverge from external reality. The conclusion we should draw from all of this is that our experience of odorants does not simply register their chemical composition and concentration, but is rather the result of computational processes that have agendas of their own. Plausibly, the representations they deliver do not stand for objective properties of any kind, but rather for viewpoint-dependent properties that are functions of functions . . . of functions of inputs. Thus, even though the olfactory system seems not to employ substantial constancy transformations, and it would therefore be inappropriate to say that it represents Thouless properties (where Thouless properties are understood strictly as the results of partial constancy transformations), it is nonetheless true that there is a significant gap between appearance and reality in the case of olfaction, and that this gap can be best explained by a model of olfactory experience that shares a basic structure with the models that plausibly explain the

experience associated with the perceptual systems that do employ constancies, vision, audition, and touch.

I have so far been proceeding under the simplifying assumption that the objects of olfaction are concentrations of molecules in the nostrils of perceivers. This is a comparatively narrow construal of the objects; but, as we have been noticing, olfaction is governed by an appreciable appearance/reality distinction even on this construal. I now add that there is a more inclusive view of the objects of olfactory awareness on which olfaction is governed by an even more appreciable appearance/reality distinction. On this broader construal, the objects of olfaction include concentrations of molecules in nostrils, but also include the larger clouds of those molecules that extend outward into space all the way to the source that is emitting them (or, in the case of mixtures, the point in space where there is a confluence of molecules from different sources). Call these more inclusive collections *odors*, and suppose that odors are the objects of olfaction. Now it is a truth about odors that they have density gradients reflecting different degrees of dispersal at different distances from their sources. It follows from this that the concentration of an odor in an agent's nostrils may be quite different from both the average concentration and the concentration at the source, which I take to be the two quantities that it is most natural to invoke in assigning a degree of concentration to the whole odor. Further, once we have seen this, we can also see that the experiences we have as a result of contact with the portions of an odor in our nostrils may be a poor guide to the strength of the whole odor. They may be rather insipid as compared to the strength of the odor itself.

To summarize, there are two ways of thinking of the objects of olfactory awareness—there is a narrow construal and a construal that is much more inclusive. On the narrow construal, the distance between olfactory appearance and olfactory reality is determined by a range of internal factors, each of them a mechanism that computes a different function of the input. On the inclusive construal, the distance between appearance and reality is determined by all of these factors, but also by the size of the gap between the strength of the proximal stimulus and the strength of the "distal" stimulus, which includes the proximal stimulus as a part.

X. Conclusion

This chapter has been partly concerned to buttress the conclusions about the contents of visual representations that we arrived at in Chapter 2, and partly concerned to determine whether, and if so to what extent, those conclusions can be extended to other perceptual modalities. It began with several arguments showing that the Thouless property hypothesis is preferable to alternative explanations of the forms of perceptual relativity that result from external conditions like distance and lighting. The next step was to establish that the hypothesis helps us to

understand the forms of perceptual relativity that are due, respectively, to characteristics of peripheral sense organs and to internal factors such as attention and adaptation. The third step was an argument that it makes sense to think of Thouless properties as the representata of experiential representations even though they are not causes of those representations. Finally, we considered what the appearance/reality distinction comes to in the case of hearing, the case of touch, and the case of smell. We found that in the first two cases, the Thouless property hypothesis could be usefully extended to explain the objects of immediate experiential awareness. This probably isn't true in the case of smell, because there is little evidence of constancy transformations in that modality, but smell is shot through with appearance/reality contrasts of other kinds, making it plausible that the immediate objects of olfactory experiences are viewpoint-dependent properties.

4
Perceptual Awareness of Particulars

I. Introduction

This chapter is principally concerned with two problems involving perceptual awareness of particular objects—as opposed to perceptual awareness of properties and relations.

The first problem has to do with the role of representation in awareness of particulars. Philosophers on one side of this issue maintain that particular objects figure in the representational contents of perceptual experiences, and that this representational relation fully explains how awareness of particulars is achieved—a subject S is aware of a particular X just in case S is perceptually representing X. Philosophers on the other side deny that particulars figure in contents. They maintain instead that perceptual experiences represent properties and relations as instantiated without specifying the instantiating objects. That is to say, on this view, perceptual experiences represent environmental situations as satisfying purely general, existential conditions involving properties and relations. The view acknowledges that there is such a thing as perceptual awareness of particulars, and seeks to explain what it involves; but it denies that such awareness involves participation in the representational contents of perceptual states.

Second, philosophers are interested in questions about the metaphysical nature of objects of perceptual awareness. Are objects of awareness members of natural and artifactual kinds (that is, complex entities like trees, ducks, and bicycles), or are they more basic entities with simpler natures, as various vision scientists and developmental psychologists seem to maintain—that is, objects defined by such simple conditions as boundedness, cohesion of parts, and spatio-temporal continuity? And if we are in some sense perceptually aware of entities of both of these kinds, what is the relationship between these two forms of awareness?

I will describe these problems in greater detail and propose solutions.

II. Particularism vs Existentialism

Representationalism is the view that perceptual states have representational contents, and that such contents are constitutively involved in perceptual awareness of objects and properties in the external world. I will assume in the present

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0004

chapter that some version of representationalism is correct. There are a number of strong reasons for making this assumption, but I will not review the full case for representationalism here. Instead I will just remind the reader that a robustly representationalist view of perceptual states is presupposed in most branches of contemporary perception science. Because of this, the success of perception science speaks loudly in favor of representationalism. (For evidence of a pervasive commitment to representations in vision science, see textbooks like Palmer 1999, Frisby and Stone 2010, and Yantis 2014. For philosophical discussion of the evidence, and additional supporting arguments, see Chapter 1 and Chapter 5.)

I will also assume that the representational contents of perceptual experiences are states of affairs, and that states of affairs can be identified with classes of possible worlds. This last assumption, about the metaphysical nature of states of affairs, is not essential. I make it only in the interest of definiteness; what I will have to say will not depend on it. (It is, however, defended in Chapter 8.) All that matters here is that states of affairs are constructed in some way from external objects and properties—that is, from environmental entities whose nature and distribution can influence the success or failure of human actions. It is quite plausible that it is the function of perception to inform us as to how things stand with such entities.

Particularism and *existentialism* are two representationalist views about the nature of perceptual contents. In my opinion, the former is false and the latter is true. I will be making a case for this assessment in the present chapter.

According to particularism, particular objects are represented by perceptual experiences, where this is understood to mean that the representational contents of experience have particular objects as constituents. Consider, for example, the visual experience that I am having at this moment—an experience of a tree out in the yard, on the other side of this window. A particularist would maintain that my experience has a representational content that includes the tree. By the same token, a particularist would maintain that if I were looking instead at a tree that was merely a twin of this particular tree, the content of my experience would be different. The experience would represent the same properties, but it would attribute those properties to a different object (Braun 1993, Soteriou 2000, Tye 2009, Tye 2014, Schellenberg 2010, Schellenberg 2016, Schellenberg 2018).

Existentialism is also concerned with the contents of perceptual experiences, but it makes a quite different claim about their nature. More specifically, it denies that particular objects figure in perceptual contents. Instead, it maintains, perceptual contents are general, in the sense that they can be fully expressed by existentially quantified propositions—specifically, by propositions of the form *there is a (single) object with such and such perceptible qualities at such and such an egocentric location*. In other words, perceptual experience informs me that certain properties and relations are instantiated in my current perceptual environment, but it carries no information about the specific identities of the instantiating entities

(Lewis 1980, McGinn 1982, Davies 1992, Byrne 2001, Hill 2009, Batty 2010, Hill 2019).

Existentialism does not deny that there is such a thing as perceptual awareness of particular objects. On the contrary, as we will see, it affirms that such awareness exists and occurs constantly. It just maintains that awareness of particular objects involves more than representation. In more positive terms, according to existentialism, a perceptual experience E provides an observer with awareness of an object O if O approximately satisfies the existential proposition that expresses the content of E, and O also stands in a certain privileged causal relationship with E. I will make an attempt to spell out what approximate satisfaction and privileged causation come to in a later section.

I will begin by giving an argument for existentialism. The next step will be to criticize the arguments for particularism. After that I will move on to present a positive account of awareness of particulars that is compatible with existentialism.

III. An Argument for Existentialism

It is a signal advantage of existentialism that it assigns a full share of representational content to hallucinations. This is desirable because hallucinations play a causal/explanatory role, and also a normative role, that depends on hallucinations having a full share of content—that is, fully determinate veridicality conditions. Particularism cannot manage this.

Consider Macbeth's hallucination. According to folk psychology, it is entirely appropriate to describe Macbeth's experience by saying that it seems to him visually that there is a dagger before him. As we are assuming representationalism, we can express this by saying that Macbeth has a visual experience with the representational content *there is a dagger before me.* Now because the experience has this content, it can play a number of causal/explanatory and normative roles. It can explain why the experience causes Macbeth to form the belief that there is a dagger before him, and why the experience causes Macbeth to extend his hand and make a grasping movement. It can also explain Macbeth's surprise and frustration when his hand closes on thin air. Moreover, it can explain how the experience justifies Macbeth's belief, and how it serves as a reason for performing the action. The thing to observe here is that the experience could not play these causal/explanatory and normative roles if it had no content, or if its content was different in kind. To be sure, it may well be true that the perceptual state that possesses the given content also has a formal or syntactic dimension, and it may also be true that the formal or syntactic properties of the state can explain why Macbeth forms a belief with certain formal or syntactic properties. But it isn't possible to explain

why Macbeth forms a belief with the content *there is a dagger before me* by appealing only to the formal or syntactic properties of the given perceptual state, nor to explain why Macbeth is justified in forming that belief. Nor can the formal properties of the state explain why Macbeth performs actions with a certain goal or aim.

It is a virtue of existentialism that it squares with this picture. Like folk psychology, it maintains that Macbeth's experience has the content *there is a dagger before me*. (I am abbreviating. A sophisticated version of existentialism would probably reject the idea that a perceptual experience can represent a high-level property like *dagger*. On such a view, Macbeth's experience would have a more basic content involving appearance properties of the dagger. In this section and at several other points in this chapter, I will prescind from the doctrine of appearance properties in discussing existentialism.) Further, because existentialism assigns this content to the experience, it honors all of the foregoing causal and normative facts. On the other hand, because particularism cannot assign a full content to Macbeth's experience, there being no particular dagger in Macbeth's vicinity that could figure in a content, it cannot make sense of the facts.

Macbeth's hallucination is easily accommodated by the existentialist explanatory framework because it does not purport to be concerned with any particular dagger. That is, while Macbeth represents the presence of a dagger, there is no particular dagger that he represents as present. Similarly, Huck Finn's alcoholic father seems to see pink elephants, but they are just generic elephants, not particular elephants that he might have admired in a zoo. It should be acknowledged, however, that there are also hallucinations that are in some sense concerned with particular objects. Milton dreamed that he saw his late espoused saint (one of his dead wives—arguably Katherine Woodcock). A boy obsessed with the Marvel Superheroes might have a vivid dream of Spiderman. What can an existentialist say about such cases? I suggest that labels like "late espoused saint," "wife," and "Katherine Woodcock" are attached to the purely perceptual components of hallucinatory experiences but are not constituents of them. It is a natural view that the labels are conceptual or linguistic in nature, and that they are confined to perceptual *judgments* that are caused by purely perceptual states. The visual system isn't sensitive to properties like *being a late espoused saint* or *being Katherine Woodcock*, though it is of course sensitive to a complex combination of perceptible features that happen to have been exemplified by Katherine Woodcock's face (and would also have been exemplified by any twin of Woodcock) (Chang and Tsao 2017). On this view, the purely perceptual content of Milton's dream was something of the form *there exists someone at such and such distance and in such and such direction with such and such facial features arranged in such and such an expression.* Because the dream was vivid and because the facial features in question were the same as the ones represented by Milton's vivid memories of his wife,

it occasioned perceptual judgments constructed from concepts that applied to his wife, either in reality or in Milton's imagination. (She sounds a bit too good to be true!) In general, it is plausible that when hallucinatory experiences come labelled with proper names or descriptions in terms of high-level properties, we should understand the labels to be components of perceptual judgments that accompany the experiences.

Assuming all of this is right, we are in a position to consider the following argument for existentialism:

First premise: The representational contents of hallucinatory perceptual experiences can be fully expressed by existential propositions.

Second premise: Veridical experiences have representational contents of the same kind as hallucinations.

Conclusion: All perceptual experiences have representational contents that can be fully expressed by existential propositions.

I will call this the *hallucination argument*.

The first premise of the argument receives strong support from the foregoing discussion.

As for the second premise, it is supported by the fact that we can't discriminate between (truly compelling) hallucinatory states and their veridical counterparts on the basis of first person access, causal/explanatory powers, or normative role. They are introspectively the same; they have the same mental and behavioral consequences; and they have the same ability to justify beliefs and provide reasons for actions. These are excellent reasons for regarding hallucinatory and veridical experiences as states of the same kind, and therefore as states that have representational contents of the same kind. (Drug induced hallucinations can be compelling, but the prize for compellingness probably belongs to the hallucinations of patients with Charles Bonnet syndrome. Here is a description: "[CBS] hallucinations are seen in perfect detail. CBS patients invariably learn that such experiences are hallucinations; yet they can be so compelling that patients are often left uncertain of whether a given object is real or not. Many test the reality of such experiences, for example, by reaching out to touch them or assessing their plausibility" (ffytche 2013, 50).)

I will call this the *argument from shared contents*. I will consider four ways of attempting to resist it.

First, one might try to resist the argument by claiming that our reasons to regard hallucinatory and veridical states as states of the same kind really only motivate the view that they are the same in point of intrinsic or formal properties. It is part of this view that the causal/explanatory powers of a perceptual state supervene on its intrinsic properties, not on its representational content.

Accordingly, there is room to suppose that if hallucinatory states explain the same introspective states as their veridical counterparts, and explain the same beliefs and actions, this is because they resemble the latter states intrinsically. We need not also suppose that they resemble the latter states in contents.

This gambit fails because, as we noticed earlier, the intrinsic or formal nature of a perceptual state can only explain the intrinsic or formal natures of the beliefs it causes. It cannot explain the beliefs qua bearers of representational content. Equally, while the intrinsic nature of an experience can explain behavioral movements, it cannot explain the goal-directedness of actions, the fact that they are aimed at achieving some end. Nor, for the same reason, can it contribute to explanations of the success or failure of actions. If Macbeth's hallucinatory experience had no representational content, but only an intrinsic nature, we could not explain his failure to grasp a dagger by saying that the experience is deceptive or misleading. Moreover, in addition to failing to do justice to the explanatory power of perceptual experiences, their intrinsic natures fail to account for their normative properties. A set of sentences cannot provide a reason to accept another sentence if one of the original sentences is meaningless. By the same token, when perceptual states are construed simply as loci of intrinsic or formal properties, they lose their ability to serve as reasons, whether for beliefs or for actions.

To summarize: If we are to continue our practice of viewing hallucinations as having causal/explanatory powers like those of veridical perceptual states, and as having the same normative powers, we must recognize that the former have representational contents.

This brings us to a second way of attempting to resist the argument for shared contents. It is sometimes maintained that we can account for the roles that hallucinations play in causation and justification without supposing that hallucinations can have exactly the same representational contents as veridical states (Tye 2014, Schellenberg 2018). We can explain these roles adequately, it is maintained, by supposing that hallucinations have only partial contents—that is, contents that are incomplete. On such a view, if a hallucination is indistinguishable from a veridical experience we might express by saying "that is a dagger," its content would contain the property of being a dagger, but there would be no particular object in the content by which the property was instantiated.

This proposal is more complex than existentialist proposals, which claim that the contents of hallucinations and veridical experiences are the same. More importantly, though, it fails because a state with a partial content could not play the same roles in causation and justification as a state with a full share of content. A perceptual state with a partial content could not explain or justify a belief with a full content. At best, it could only provide a partial explanation or a partial justification—that is, an explanation or justification of some part or aspect of the content of the belief. But it is clear that hallucinations can lead agents to form

beliefs about their environments that are false, and therefore furnished with a full share of content. Similarly, a perceptual state with a partial content could not provide a full explanation or a full reason for an action. When we act on the basis of perception, it is because our perceptual experience represents the world as being in a certain way. But for the world to be a certain way is for there to be particulars that exemplify certain properties and relations. Hence, a perceptual state that explains or justifies an action must represent properties and relations as exemplified, either by attributing the properties and relations to specific particulars or by representing them as exemplified by some particulars or other. That is to say, the explanatory and justificatory powers of perceptual states depend on their representing full states of affairs, not just free-floating properties and relations.

Third, one might attempt to block the argument for shared contents by allowing that hallucinatory states and veridical states both have full contents, but maintaining that the contents of the former are different in kind from the contents of the latter. More specifically, it might be maintained that while the former have existential contents, the latter have contents that involve specific particulars.

There are two problems with this response. First, like the "partial content" proposal, it is more complex than existentialism, which offers a uniform account of perceptual contents. And, second, the response would saddle us with the impossible task of explaining how and why evolutionary processes or perceptual learning arrange for perceptual states to have existential contents in hallucinatory contexts. Let R be a representation that is tokened in hallucinatory contexts. By hypothesis R has an existential content involving one or more external properties—that is, a content like *there is a dagger before me.* How could it have acquired this content? Normally, contents are acquired by virtue of causal or probabilistic relations between representations and external states of affairs, but, by definition of "hallucination," there is no such relation linking R to anything external. R is produced entirely by internal causes, and is only linked probabilistically to internal phenomena. Accordingly, it *could not* have acquired an external content like *there is a dagger before me.*

There is no similar problem confronting existentialism, for existentialism maintains that hallucinatory representations are identical, both intrinsically and in point of content, with the representations that figure in veridical perception. It follows that existentialism need not posit a separate process by which hallucinatory representations acquire their contents. It can comfortably maintain that the representations acquire their contents in veridical contexts and continue to possess them when they are generated by internal causes.

Fourth, it might be maintained that veridical and hallucinatory experiences share what David Kaplan calls a *character*. The idea of a character originated as a solution to a problem in the philosophy of language. Consider the indexical sentence "I am happy." It is natural to suppose that this sentence has different

meanings in different contexts: in a context in which Arif is the speaker, it represents the state of affairs *Arif is happy*; in a context in which Bosum is the speaker, it represents the state of affairs *Bosum is happy*; and so on. But although the sentence represents different states of affairs in different contexts, and in that sense has different meanings in different contexts, there is also a sense in which it has the same meaning in every context. When uttered in any context *C*, the sentence attributes the property of being happy to the person who is the speaker in C. Kaplan sought to model this second kind of meaning, which is context-independent, by saying that in addition to representing many different states of affairs in different contexts, the sentence also has a character, which he took to be a partial function *f* that has two properties: (i) *f* maps any context of the sentence in which there is a speaker X onto the state of affairs *X is happy*; and (ii) *f* is undefined for any context in which there is no speaker (Kaplan 1989). Kaplan viewed this character as a semantic property of the sentence that it possesses relative to any context, independently of whether the sentence is true or false in that context, and even independently of whether there is a speaker in that context. Now Michael Tye has suggested that Kaplan's idea might be fruitfully transposed from the philosophy of language to the philosophy of perception (Tye 2009 and Tye 2014). Let E be a type of perceptual experience that we might express by the sentence "That cube is red." Like any visual experience, E can occur in contexts in which a perceiver is viewing an actual cube and also in contexts in which E is hallucinatory. It makes sense to say that E represents a particular cube in any context of the first kind, but this would clearly be the wrong thing to say in contexts of the second kind. Yet, despite this difference in the semantic properties of E across contexts of these two kinds, we might also want to say that there is a sense in which E has the same semantic nature relative to all contexts. On Tye's proposal, we should model this second, context-independent semantic value of E by supposing that E is associated with a character-like function *f* with the following binary nature: (i) for any context C in which a subject is actually perceiving a cube X, *f* maps C onto the state of affairs *X is a cube and X is red*; and (ii) *f* is undefined for any context in which a subject is not perceiving a cube. *f* is a semantic value of E that it possesses relative to all contexts. Alternatively, we can say that it is a semantic value that is shared by all tokens of E. The core of Tye's idea is that we might be able to appeal to this shared semantic property to explain why all tokens of E have the same causal and normative roles.

This proposal has a certain initial appeal, but reflection shows that it fails for the same reason that the partial content proposal fails: like partial contents, characters are not states of affairs, so they cannot figure in explanations of goals, nor provide reasons for belief and other attitudes. Recall that characters, as I have defined them, are partial functions of contexts of experiences. Something that belongs to this ontological category is worlds apart from the states of affairs that

are under consideration here. Unlike states of affairs, characters cannot be said to be actual (realized) or non-actual (unrealized). They have no bearing on how things are actually arranged in the world.

It appears, then, that the argument from shared contents can withstand criticism from particularists. Together with the discussion at the beginning of this section, that argument provides strong support for the hallucination argument for existentialism. It seems fair to conclude that at this point in the race, existentialism is well ahead of particularism.

The case for existentialism I have been developing is based entirely on considerations having to do with hallucinations, but there are other ways of arguing for the view. I will call attention to one of these alternatives at the end of the next section.

IV. Arguments for Particularism

Because it insists that all veridical perceptual contents involve particulars, particularism has trouble accommodating the contents of hallucinations like Macbeth's experience as of a dagger. In Macbeth's case, there is simply no plausible candidate for the role of an embedded particular. Accordingly, particularists are forced by hallucinations to tell a more complex story than existentialists about the nature of perceptual contents. More specifically, they must tell a disjunctive story, maintaining that hallucinations differ from veridical experiences in that they have no content, or that they differ in that their contents are partial, or that they differ in that they have full contents that are different in kind than veridical contents. As we have seen, in addition to being more complex, these stories are extremely implausible.

Still, there are considerations that seem, prima facie at least, to provide support for particularism. I will review three of them and offer responses.

1. According to some philosophers, when one is undergoing a veridical perceptual experience, one allegedly has a vivid impression of particularity—an impression that seems to be grounded in the phenomenology that is part of the experience. When, for example, one focuses on the face of a friend, one has a compelling impression that one is experiencing a particular human being. It has been maintained that this sort of *phenomenological particularity* counts heavily in favor of particularism.

In response, I suggest that what is sometimes described as an impression of particularity is best understood as an impression of uniqueness. It is quite true that perceptual experience represents clusters of properties as uniquely coinstantiated. Thus, when I have a perceptual experience, I represent the environment as containing an object that is at a certain specific egocentric distance from me and

in a certain specific egocentric direction. These spatial properties could not be possessed by two objects (except in very special circumstances that I will mention in a moment). It is easily seen, however, that this sort of uniqueness is fully compatible with existentialism. Existentialism comfortably accommodates perceptual contents of the form *there exists an object at such and such distance and in such and such direction and with such and such other properties (size, shape, etc.)*. (The very special circumstances I mentioned a moment ago involve objects that are exhaustively constituted by other objects that begin and end at the same time. An example due to Allan Gibbard is a statue (*Goliath*) that is constituted by a lump of clay (Lumpl). Both *Goliath* and Lumpl come into existence when lumps of clay constituting the head and torso of a statue are joined to form a new statue (*Goliath*) and a new lump (Lumpl). They also end at the same time, due to someone's pulling the head and the torso apart. In this case, Goliath and Lumpl have the same polar coordinates, but they are different objects. The viewer has an impression of uniqueness. The impression is, however, illusory (Gibbard 1975).)

A particularist might respond by claiming that impressions of particularity are not reducible to impressions of uniqueness; but to say that is to assume the burden of explaining how exactly impressions of particularity differ from impressions of uniqueness, and one could be forgiven for being skeptical about such an endeavor. It is not as if perceptual experience involves two types of phenomenology—one type that consists of uniqueness-entailing properties, and a second type that is indicative of particularity. What could that even mean? If I were asked to withdraw my attention from such features of an object as apparent size and apparent location, and to focus instead on the aspect of phenomenology that is indicative of particularity, I wouldn't know how to proceed. Or is it the fact that phenomenal properties are bound together that is supposed to be indicative of particularity? That seems unlikely, since binding is fully accommodated by existentialism.

Why are philosophers occasionally moved to claim that there are impressions of particularity? Part of the explanation is that it is possible to confuse uniqueness with particularity, but there is also another reason. As we noticed in connection with Milton's dream of his long-departed wife, perceptual experiences are often bound to singular concepts so closely that introspection cannot discriminate between conceptual content and content that is purely experiential. What was I experiencing a moment ago? The natural answer is that I was experiencing Lee, but it is plausible that my experience counted as an experience of Lee because it automatically triggered my concept of Lee. Equally, an experience attesting to the existence before me of a woman wearing a red dress might automatically trigger a descriptive concept, such as *the woman in the red dress*, or a demonstrative concept, such as *that woman in red.*

2. It is sometimes maintained that since the contents of concepts derive to a large extent from the contents of perceptual experiences, we would not be able to account for the particularity of conceptual content unless perceptual content was also characterized by particularity. How could I use the concept of Lee or the concept of this laptop to single out particular objects if my use of the concepts wasn't grounded in experience of those objects?

The answer to this question is that it isn't necessary to represent a particular object perceptually in order to be perceptually aware of it. Earlier I noted that existentialists are no less committed than particularists to the existence and prevalence of perceptual awareness of individual objects. They just have a radically different conception of what this form of awareness consists in. As we then observed, the idea is that when a subject is perceptually aware of an object O via a perceptual representation R, O must approximately satisfy the existential content of R, and O must stand in an appropriate causal relation to R. I will return to this proposal later on.

3. A third argument for particularism maintains that it is uniquely well qualified to explain our ability to keep track of objects across time.

If perceptual contents were purely existential, it might seem, then the content of each member of a series of experiences would be independent of the contents of the other members, just as the content of "Someone came into the room at time T" is independent of the content of "Someone turned on the light at time T + 1." As is obvious, these two sentences do not entail "Someone came into the room at time T and turned on the light at T + 1." Accordingly, they cannot justify one in accepting the third sentence, or provide a ground for acting as if the third sentence is true. By the same token, if each member of a series of experiences had a separate existential content, the series would not be able to justify the belief that a single object had serially possessed the properties with which the members of the series were concerned, nor could the series provide a reason for acting as if there were such an object. (That is, a series of experiences with the existential contents (Ex) F_1x, (Ex) F_2x, . . . , (Ex) F_nx could not justify a belief with the content (Ex) (F_1x & F_2x & . . . & F_nx), nor could the experiences provide a reason for acting as if that content were veridical.)

In view of these considerations, it can seem that existentialism lacks the resources to account for the fact that we can track objects as they proceed through time. On the other hand, particularism is well equipped to explain diachronic tracking. In standard versions, it maintains that a series of experiences can share a singular referring element, and that this element can continue to pick out the same object in each member of the series.

According to Zenon Pylyshyn, for example, if we are to explain tracking of objects across time, we must suppose that the visual system employs special referential devices that he calls "visual indexes" (Pylshyn 2003, Pylyshyn 2007).

Pylyshyn describes these indexes as counterparts of linguistic demonstratives like "this" and "that." Like such demonstratives, visual indexes pick objects out directly, without invoking properties of any kind—not even properties having to do with locations in space. To borrow a phrase from the philosophy of language, they are *purely referential*. When asked how exactly these indexes manage to acquire reference, Pylyshyn answers that the connection between visual indexes and their referents is determined by a purely causal relationship: indexes are somehow causally triggered by objects, and once these causal connections have been established, they can persist through changes of various kinds. Pylyshyn accounts for diachronic tracking by claiming, first, that a single index can be shared by all of the members of a series of experiences, and, second, that the reference of an index can be preserved across the members of a series, even if the series is quite long. Pylyshyn allows that representations of properties can become bound to visual indexes. His point is just that the reference of an index isn't determined by the representations of properties that become attached to it.

Pylyshyn maintains that visual indexes emerge in early vision. Particularism can also be developed by appealing to *object files*—theoretical entities that are very similar to visual indexes except that they are usually thought to be associated with mid-level vision (Kahneman and Treisman 1984, Kahneman et al. 1992, Palmer 1999, Green and Quilty-Dunn 2017).

There are three good reasons to doubt purely causal accounts of how reference is determined. First, as Ned Block has pointed out, early vision has the task of distinguishing figure from ground, and that task requires information about borders of the figure. In brief, it requires information about shape. Here is Block explaining this point (Block 2022, chapter 3):

> [T]he figure/ground structure of much perception requires a difference between what is attributed to the figure and to the ground. Perception of the figure and the ground requires not only distinguishing figure from ground but also perceiving the figureas having a particular shape—and that requires attribution. Attribution of shapes occurs "automatically and obligatorily" (Baker & Kellman, 2018, 1295), even when there is no task that requires such attribution.

Thus, according to Block, in order to introduce or maintain a referential relation between a visual index and an object in the world, it would be necessary to discriminate the object from its background, and, to achieve this, it would be necessary to represent the shape of the object and attribute that shape to the object. It follows that representation of properties is a necessary condition of representation of objects.

A second reason for doubting purely causal accounts of reference to particulars has to do with tracking objects across time. As is shown by experimental work of

various kinds (e.g., Pylyshyn's own studies of multiple object tracking and the studies of children's awareness of object permanence by Rene Baillargeon (1985) and others), it is possible to keep track of objects visually even when they are hidden behind occluders. But, of course, when objects are hidden behind occluders, they are no longer capable of causally entraining referring devices. Accordingly, diachronic tracking cannot depend purely on causal relations between such devices and objects. Rather, it must involve the use of descriptive information about the properties and locations of objects. (When an object O passes behind an occluder, and an object O* emerges on the other side, O* must be on more or less the same trajectory as O in order for subjects to treat the objects as the same. It is not entirely clear what constitutes sameness of trajectory; but even so, it is apparent that subjects could not determine whether O and O* are on the same trajectory unless they register and remember the location of O as it is passing behind the occluder, and in some way compare it to the location where O* appears.)

This brings us to the third reason for rejecting Pylyshyn's view: reflection on the causal relationships between objects in the world and visual indexes shows that Pylyshyn has an overly simple view of what those relationships must involve. Thus, Pylyshyn thinks that objects causally entrain visual indexes as they move, without making use of any representations of the objects' properties. In short, his indexes are pure demonstratives. But if different objects in a multiple object tracking (MOT) experiment entrain different indexes, they must do so in virtue of having different properties; and, given the set up, locations are the only differentiating properties that could do the requisite causal work. (All the objects in a classic MOT set up are identical in point of shape, color, size, and other intrinsic properties. There is a more detailed description of MOT experiments in Section VI.) Further, the indexes must have distinguishing properties that correspond to these locations, for otherwise it would be impossible for a particular object to entrain a particular index. So, at any one time, each of the objects being tracked has a locational property L that distinguishes it from all of the other objects being tracked, and the corresponding index has a property P that distinguishes it from all of the other indexes that are being deployed on that occasion. Moreover, it must be true that at any time T, the given object causally entrains the given index in virtue of possessing L at T, and that the index is entrained by that object in virtue of possessing P. This leads naturally to the question, "Why not say that P represents the object as being at L at T?" As far as I can tell, Pylyshyn has no answer to this question. But current theories of perceptual representation suggest that we should see a representational relationship here (Burge 2010, Neander 2017, Shea 2018).

It appears, then, that Pylyshyn's views about diachronic perceptual tracking are misguided. But this is only part of what needs to be said in defense of existentialism. It is necessary to add that existentialism is perfectly capable of accounting for

diachronic tracking. Here is one form that such an explanation could take: Suppose that synchronic perceptual representations have a formal structure that can be represented as follows: "item X has the properties F_1, F_2, . . . F_j at time T_n." That is, suppose that from a formal or syntactic point of view, perceptual representations consist of free variables and representations of properties. Suppose also that temporally extended series of such representations can share a variable, and that when a variable occurs in all of the members of a gapless series of representations of this sort, it has the effect of grouping the members of the series together, producing temporally extended but unified representations that are counterparts of this schema:

(*) Item X has the properties F_1, F_2, . . . , F_j at time T_1 and item X has the properties G_1, G_2, . . . , G_n at time T_2 and . . . and item X has the properties H_1, H_2, . . . , H_p at time T_n

Finally, we can suppose that the perceptual counterparts of (*) are bound by devices that are in effect quantifiers—quantifiers whose scope expands as time passes and new conjuncts are added. These devices owe their status as quantifiers to the fact that perceptual states containing them have the function of encoding information about existential states of affairs (that is, classes of worlds whose specification requires an existential quantifier or some equivalent set of operators). We are justified in positing them by the need to account for the explanatory and justificatory powers of perceptual states. When they are affixed to representations of form (*), we can think of the resulting complex representations as object files. (The scope of a quantifier ends when a conjunction stops growing.)

What explains the fact that the members of a temporally extended series are grouped together as the series expands? The answer lies with relationships between the individual synchronic representations that serve as members of the series. To illustrate, the locations attributed by successive synchronic representations must be such as to describe a continuous path, except for interruptions by occluders, in which case the hidden parts of the trajectory that is attributed to an object must be smooth. Moreover, the later members of the series must attribute boundaries to an object that can be obtained from earlier boundaries by continuous stretching and shrinking. Another condition is that patterns of connectedness within those boundaries must be preserved. These conditions are all well motivated. Indeed, reflection shows that they would have to be honored in any theory that posited singular representations. How could a visual demonstrative continue to refer to the same object across time if there were sharp discontinuities in the spatio-temporal path that was attributed to the object?

Incidentally, this line of thought challenges an argument in Section I of Green 2017b. According to Green, we would not be able to track objects across change if perceptual representations had descriptive content, the alleged problem being

that the representations could not change descriptive contents across time without losing their identities. This overlooks the possibility of temporally extended representations whose stages are held together by the relations specified in the foregoing paragraph.

To summarize: Although there are reasons for thinking that particularism is true, these reasons either evaporate upon close examination or turn out to be misleading, in the sense that they aren't reasons for *preferring* particularism to existentialism.

At the end of the previous section I promised to sketch a reason for embracing existentialism that is independent of the hallucination argument. We are now in a position to formulate a second reason. In the earlier paragraphs of this section, we reviewed attempts to justify particularism at the expense of existentialism. What we found is that attempts to motivate particularism fail: in one case, the alleged data to which an attempt appeals turn out to dissipate upon examination; and in others the data to which the attempts appeal can be comfortably accommodated by existentialism. Now in addition to undercutting the motivation for particularism, these conclusions implicitly support existentialism. For the former view is more complex than the latter. Both views recognize that perceptual experience represents collections of properties as coinstantiated. They differ in that existentialism claims that this is the sole content that perceptual experiences possess, while particularism maintains that the contents of perceptual experiences have additional components—specific particular objects that instantiate the collections of represented properties. Since particularism makes the stronger claim, it is in need of supporting considerations beyond the considerations that provide motivation for existentialism. But, as we just observed, the attempts to find proprietary support for particularism don't work out.

V. Awareness of Particulars

I will not attempt to state a formula that applies to all of the forms of perceptual awareness of particular objects. Awareness of objects is different in the case of vision than in the case of hearing, and different in the case of hearing than in the case of olfaction. In the interests of definiteness, I will focus on visual perception. This will suffice for present purposes. Another limitation is that I will not attempt to state conditions that are both necessary and sufficient for visual awareness. A sufficient condition is all that I seek. Visual awareness of objects comes to the same thing as seeing objects, and our commonsense concept of seeing is quite vague—so vague that it is often unclear how exactly it applies in penumbral cases, such as cases in which the locations of objects are badly misrepresented. Moreover, the concept is extremely labile, applying to what seems to be an

impossibly broad range of cases. It is possible to see fires, rainbows, shadows, holograms, mirror images, virtual reality images, and patches of sky. In view of this range, it seems unlikely that there is a single, unified condition that is both necessary and sufficient for seeing. Perhaps seeing is a family resemblance concept, or a concept that has a paradigms-and-similarity architecture. Anyway, a sufficient condition is all that is required for present purposes, as long as it is clear that it is satisfied in a significant range of cases, and that those cases include paradigmatic situations in which we rely heavily on visual awareness in fixing the reference of demonstratives, proper names, and kind names. The goal is to show that existentialism is compatible with the indisputable fact that vision puts us in touch with particular objects in normal situations, and also with the indisputable fact that the reference of many words and concepts is inherited from visual awareness.

Here is a proposal that serves these purposes:

S is visually aware of O via a conscious perceptual experience E if
(i) S's experience E represents that there is a (single) object with such and such perceptible qualities in such and such a location,
(ii) O is causally responsible for E, and
(iii) O approximately satisfies the representational content of E, in the sense that it possesses some of the properties that E represents as instantiated, and possesses more of those properties than any other cause of E.

All three of these conditions require explanation and defense; but, as noted at the outset, I am assuming in this chapter that perceptual experiences have representational contents. I think we can also now assume that it makes sense to claim that those contents can be expressed by existentially quantified propositions. Relative to these assumptions, condition (i) is reasonably straightforward. Accordingly, we can focus here on conditions (ii) and (iii).

Suppose that E is a perceptual experience and that R1, R2, . . . , Rj are parts or aspects of E that have representational content. (ii) is meant to imply that if O is the object of perceptual awareness in this case, O must possess certain properties Q1, Q2, . . . , Qj that are causally responsible for R1, R2, . . . , Rj. Moreover, it is meant to imply that the causal chain running from Q1, Q2, . . . , Qj to R1, R2, . . . , Rj begins with light being reflected by O. This reflected light must be stamped with the structural signatures that are characteristic of Q1, Q2, . . . , Qj, and upon arrival at the eye of our subject S, it must trigger a chain of processing mechanisms that have the function of producing R1, R2, . . . , Rj in response to those signatures.

(ii) blocks a familiar objection to existentialist theories of perceptual content—the objection that they cannot discriminate between genuine awareness of objects and cases of veridical hallucination, where a veridical hallucination is a case in

which a subject hallucinates that certain properties are jointly instantiated in the current environment and it is in fact true, though only serendipitously, that there is an object with those properties in the environment. Cases of this sort are counterexamples to theories that seek to analyze awareness of objects purely in terms of *correspondence* relations between represented properties and properties that are actually possessed by objects, but they have no bearing on theories that invoke *causal* relations as well as correspondence relations.

To conclude this discussion of clause (ii), I note that while the properties Q1, Q2, . . . , Qj are the ones that are causally responsible for the tokening of R1, R2, . . . , Rj, they are not the properties that are represented by R1, R2, . . . , Rj. Q1, Q2, . . . , Qj are properties that can impose structure on reflected light. Hence they are objective properties like sizes, shapes, and textures. According to the doctrine elaborated in Chapters 2 and 3, the properties represented by R1, R2, . . . , Rj are different. They are appearance properties that reflect the influence of internal computations. They must, however, include properties that are actually possessed by the object O, if condition (ii) is to be satisfied. (As the reader may recall, this distinction between the properties that are causal triggers of R1, R2, . . . , Rj and the properties that are represented by R1, R2, . . . , Rj is discussed at length in Section IV of Chapter 3.)

What is the rationale for condition (iii) in the definition of visual awareness? Why don't (i) and (ii) do suffice? (iii) is needed because the properties that figure in condition (ii)—the properties that serve as the distal causal triggers for experience E—do not entail that O has the appearance properties that E represents. The properties that serve as the distal causal triggers for E are objective properties like objective sizes and shapes. The appearance properties that E represents do not depend only on distal objective properties of objects, but on internal computations as well. Accordingly, since it is an important intuitive constraint on objects of perception that they possess at least some of the properties that experiences attribute to them, there is need for a condition that speaks of attributed appearance properties as well as one that is concerned with properties that figure in the etiology of E.

VI. The Nature of Objects of Perceptual Awareness

You enter a café and scan it to see if a friend is there. As your eyes travel around the room, you see tables, chairs, cups, glasses, plates, vases of flowers, houseplants, waiters, and patrons, among other things. If someone were to say that you didn't really see these items, but inferred their existence from objects of other kinds, which you really did see, you would be baffled and annoyed. We can express this intuitive conception of the objects of visual awareness by saying that they consist principally of members of artifactual and natural kinds. In brief, the objects of ordinary perceptual awareness are principally *kind objects*. (As the café example

may suggest, I will be focusing here on the objects of visual awareness. I add that I will be prescinding in this section from awareness of exotic objects like VR images and patches of sky.)

As against this commonsense view, there is a hypothesis about the objects of visual awareness that represents them as being defined by conditions that are much simpler than the conditions that define kind objects. The evidence for this hypothesis comes from several branches of cognitive science.

Some of the evidence derives from studies of object tracking in infants—experiments in which very young subjects follow single objects as they move. It turns out that the subjects can do this successfully as long as three conditions are satisfied—cohesion, boundedness, and spatio-temporal continuity (Spelke 1990, Scholl 2007, Carey 2009, Burge 2010). As long as the conditions are satisfied, infants will follow objects visually through changes of size, shape, color, and texture—even though these qualitative changes are ones that normally betoken changes of kind affiliation. It follows that continuity of kind is not required for object tracking. Or so it can seem.

Elizabeth Spelke (1988, 1990) is the psychologist who first emphasized the roles of cohesion, boundedness, and spatio-temporal continuity in determining object identity. Accordingly, things that satisfy these criteria have come to be known as *Spelke objects*.

Another impressive body of evidence comes from studies of multiple object tracking in adults. Subjects in MOT experiments attend simultaneously to small sets of designated objects as they move around space, passing among other objects (distractors), and changing properties like size, shape, and color as they proceed. It turns out that subjects can successfully keep track of up to four such objects despite all of the changes that the objects undergo (Pylyshyn 2003, 2007, Scholl 2007). Interestingly, the conditions of success are more or less the same as the conditions that must be satisfied in order for children to keep track of moving objects (Carey and Xu 2001). In the literature, this convergence of results is sometimes expressed by saying that both of these experimental traditions indicate that Spelke objects are the fundamental objects of visual awareness.

Additional evidence for this idea comes from studies of tunneling, in which subjects track moving objects that appear to change their kind affiliations when they pass behind occluders (Flombaum et al. 2004, Scholl 2007).

It should be emphasized that Spelke objects are quite different than kind objects. Consider an evolving situation in which, first, a fertilized egg morphs into a human being, and, second, the human being morphs into a corpse. In this sequence there are three different kind objects, but only one Spelke object. For another example, consider a situation in which at time T1 there is a lump of clay, at time T2 the lump is molded into a sculpture of an alligator, and at time T3 the sculpture is transformed into a new sculpture that has the form of an owl. Here again we have three kind objects, the lump and the two sculptures, but only one Spelke object. We can summarize this situation by saying that the principles of

individuation that govern Spelke objects tend to be quite different than the principles of individuation that govern objects qua members of kinds.

A caveat: In a recent paper, E. J. Green has argued convincingly that the principles of individuation governing objects of visual awareness are even less restrictive than the ones that define Spelke objects (Green 2018). His principal point is that visual objects seem not to satisfy a strict cohesion condition, which entails connectedness of parts. Instead, it seems likely that visual objects are governed by a weaker condition requiring translational cohesion, which entails only that the parts must move together, sharing directions and velocities. This weakening might have been expected: we can after all track flocks of geese and swarms of wasps. *Pace* Green, however, I will here assume that the Spelke conditions are correct. Translational cohesion is a form of cohesion, after all, albeit a rather abstract form, so getting the right view seems to be a matter of interpreting Spelke's coherence condition properly as opposed to rejecting it wholesale. (Similarly for Spelke's boundedness condition. Boundaries need not take the form of unbroken lines.) Anyway, whether we reinterpret Spelke's hypothesis or change it more drastically, the hypothesis vividly captures the point that is essential here—the point that visual objects are defined by conditions that are much too simple and general to support a division of objects into multiple kinds.

It appears, then, that there are two conceptions of the metaphysical nature of the objects that we see and visually track through time, one that is backed by our commonsense thought and talk, and another that receives strong support from experiments in several branches of cognitive science. This divergence of opinion raises some important questions. Are the conceptions in competition, or is there some way of reconciling them, where a reconciliation might involve saying that the visual system has two parts or aspects, one that constructs and deploys representations of Spelke objects, and another that traffics in representations of kind objects? Further, should we adopt a posture of metaphysical realism toward objects of both types? If so, how are the types related? And if not, which of the two types is real?

These questions arise for both particularists and existentialists. Particularists need to specify the nature of the particular objects that figure in the representational contents of perceptual experiences, and existentialists owe us an account of the objects in the universe of discourse over which their existential quantifiers range.

VII. A Dual Systems Account of Object Awareness

It is sometimes suggested that the mind makes use of two quite different systems for processing information about objects. This idea appears at various points in

the cognitive science literature, but perhaps most notably in a well-known paper by Carey and Xu (2001). Here is a summary of their proposal:

> In adults, there is prima facie evidence that at least two distinct representational systems underlie object individuation. The first is the mid-level vision system (mid-level because it falls between low-level sensory processing and high-level placement into kind categories) that establishes object file representations, and that indexes attended objects and tracks them through time. . . . This first system . . . privileges spatiotemporal information in the service of individuation and numerical identity. Individual objects are coherent, spatially separate and separately movable, spatiotemporally continuous entities. Features such as color, shape, and texture may be bound in the representations of already individuated objects; [but] they play a secondary role in decisions about numerical identity. . . . The second system . . . is fully conceptual, drawing on kind information for decisions about individuation and numerical identity. (p. 181)

As I understand the authors, they think that both of the systems they distinguish are perceptually grounded, but they also believe that only one of them is fully perceptual. The other system makes essential use of conceptual representations.

Carey and Xu do not explicitly maintain that the systems have different ontological commitments, but that seems to be an implication of their ideas. Each system is associated with a proprietary system of principles of individuation, and these sets are quite different. The principles in the first set could not possibly be satisfied by the same entities that satisfy the principles in the second set. Accordingly, the authors are committed to saying either (i) that one or the other of the two systems is completely delusive, which entails that Mother Nature has gone to considerable trouble to endow us with a faculty that never produces veridical representations, or (ii) that there really are two metaphysically distinct types of entity with which the systems are respectively concerned, kind objects and Spelke objects. It might seem that cognitive scientists are barred from taking the latter course, on the grounds that they aren't professionally qualified to make decisions about what kinds of objects exist. Who are they to add Spelke objects to the list of entities we should believe in? But we need not think of recognizing Spelke objects as ontologically inflationary, for they can be analyzed as series of kind objects that satisfy the three Spelke criteria. For example, we can suppose that a Spelke object that starts as a fertilized egg, becomes an adult human being, and then turns into a corpse is just a series consisting of those three kind objects.

Although they are perhaps metaphysically defensible, dual systems theories face difficulties in explaining how the two systems are related. Presumably the systems do not function independently: there must be a binding relationship of some sort between the representations that they respectively produce. How else

could spatial and qualitative information be connected with information concerning kind affiliation? It would, however, be quite inappropriate to say that the binding expresses predication. If conceptual representations of kinds were bound predicatively to object files that stand for Spelke objects, the results would be category mistakes. Even if Spelke objects are analyzed as series of kind objects, the fact that they are governed by very weak principles of individuation precludes their counting as members of kinds.

VIII. A Unified Account of Object Awareness

The dual systems approach is not without attractions, but I wish to recommend a quite different set of ideas. This alternative approach maintains that objects of visual awareness are always members of kinds (hereafter *k-objects*), and are accordingly individuated by conditions that are much more complex, and also much more demanding, than the Spelke criteria. The approach is plausible for several reasons, the most important of which is that it squares with an ecological conception of the environments in which we exist. It is k-objects that answer to our biological needs, and it is k-objects that we must navigate among as we seek fulfillment of our needs. In a phrase, k-objects are loci of ecologically relevant causal powers. On the other hand, this is not generally true of Spelke objects, even if Spelke objects are analyzed as series of k-objects. The biologically relevant causal powers of Spelke objects change radically over time. A seed cannot provide much nourishment, even if it is continuous with a plant that will produce several ears of corn. And the causal powers of a living animal are vastly different than the powers of the corresponding corpse, even though the animal is continuous with the corpse. We can put this contrast by saying that human agents *need* to represent k-objects, and that there is no comparable need to represent Spelke objects. This makes it likely that k-objects, and only k-objects, are what human agents *do* represent. The same is true of members of neighboring species. (The idea that the biological needs play a substantial role in determining perceptual content is a common theme in current theories of representation. Cf. Millikan 1989a, Dretske 1995, Burge 2010, Neander 2017, and Shea 2018.)

A defender of Spelke objects might say that Spelke objects often correspond to k-objects with respect to properties like size, shape, weight, and other features that determine causal powers. For example, it might be said that a boulder tends to maintain the same features of this kind throughout its existence. It might also be pointed out that the causal powers possessed by boulders in virtue of such properties as size and weight have ecological significance—for example, in navigation. But considerations of this sort do not affect the claim I wish to make, which is that Spelke objects as a class have considerably less ecological relevance

than k-objects as a class. It follows that if we want a uniform ontology of perceptual objects, we are well advised to choose the class of k-objects.

But how can this proposal be reconciled with the evidence that the visual system tracks entities that satisfy the Spelke criteria? The answer is that tracking entities that satisfy the Spelke criteria is plausibly just a useful heuristic for tracking k-objects. This is so because in all normal cases, tracking an entity that satisfies the Spelke criteria comes to the same thing as tracking a k-object. Objects of one kind rarely morph into objects of other kinds in the course of the intervals during which they are being continuously observed. It is much, much more common for an object to remain a leaf or a bicycle or a human being during periods of perceptual tracking, no matter how extended the periods might be. Accordingly, if Mother Nature was principally concerned to equip us with the ability to track k-objects, it would be natural for Her to accomplish this by endowing us with a simple heuristic that privileges the Spelke criteria.

To be sure, this heuristic can mislead us, and typically does so in experimental settings in which transitions (or apparent transitions) are quite rapid. In these special contexts it appears to observers that there is a single object that changes its kind affiliations (Carey and Xu 2001, Flombaum et al. 2004, Scholl 2007). But to take these exceptional cases as evidence that the visual system is designed to traffic in a special class of entities—Spelke objects—would be like supposing that the tail is a special device for wagging the dog. A more appealing interpretation is that agents undergo visual illusions in these cases. One has the visual impression that identity is preserved, but there is no corresponding external fact. (Scholl (2007) makes a related suggestion.)

Note that postulating a system that has the function of tracking k-objects is quite different than postulating a system that has the function of tracking Spelke objects, even if the former is based on the Spelke criteria. A system of the former kind produces illusory representations when one k-object morphs into another, while a system of the latter kind gets things right in cases of morphing. Perhaps dual systems theorists have overlooked this distinction.

IX. Awareness of Objects Qua Members of Kinds

It is plausible, then, that the objects of visual awareness are k-objects. But there is awareness and there is awareness *as*. If we are visually aware of k-objects, is it also true that we can be aware of them *as* members of specific kinds? No. To be sure, we are clearly capable of visually discriminating the gazelles from the lions at the watering hole, and Ferraris from sixteen-wheel trucks on the highway. We are also capable of responding to stimuli differentially in ways that reflect such discriminations. But it by no means follows that we perceptually represent gazelles

as members of the kind *gazelle*, or that we perceptually represent lions as members of the kind *lion*. It is possible to explain our discriminative abilities by supposing that we simply represent gazelles and lions as possessing congeries of low-level properties, such as shape, that are strongly correlated with kind membership.

In "Is Semantics Possible?" and "The Meaning of 'Meaning,'" Putnam argued that natural and artifactual kinds are not individuated by purely perceptual criteria. Rather, their individuation may depend on such remote factors as biological ancestry, chemical microstructure, and designs prepared by an engineer (Putnam 1975). This view has since been confirmed by developmental psychologists (Keil 1989, Gelman 2003). But Putnam also maintained that we have perceptual stereotypes of natural kinds—representations of the salient observable properties of normal members of the kinds. He argued that these perceptual stereotypes provide a perceptual grounding for our concepts of kinds, and guide the use of such concepts. Prima facie, at least, these views are plausible. It is also plausible that there are stereotypes of artifactual kinds that play a similar role with respect to our concepts of artifacts.

Does science provide grounds for accepting Putnam's picture? Yes. It has long been recognized in vision science that low-level properties, and in particular shape, play a large role in categorization (Biederman and Ju 1988), but recent work has provided much stronger confirmation for Putnam's picture. Perhaps the most striking piece of evidence is a recent study by Chang and Tsao (2017) showing that all human faces can be represented by points in a fifty-dimensional space, where each dimension corresponds to the degree to which a face instantiates a certain low-level feature. Evidently the kind *human face* is strongly correlated with—indeed, is uniquely determined by—the second order property *instantiates a point in the space defined by dimensions D1, D2, . . . , D50*. Further, in addition to evidence of this kind, which comes from vision science, there is evidence for Putnam's view in the more cognitively oriented literature on categorization. There are a number of different models of how perceptually grounded categorization works, but virtually all theories incorporate the Putnamian idea that perception-based categorization involves representations of statistical relationships among the perceptible features of kinds (Murphy 2002, Machery 2009). It appears that perceptual stereotypes are all we need to explain perceptual discrimination among specific kinds and differential responses to kinds. I see no advantage in the more ambitious hypothesis that we perceptually represent members of kinds *as* members of kinds (though representational pluralism allows for the possibility that we can be said to represent objects as members of kinds in a secondary sense).

To summarize, while experiments in vision science and developmental psychology have some tendency to suggest that the objects that are tracked by the visual system are defined by very weak and highly general criteria, this view is at

odds with the idea that visual objects as such have ecological significance. Spelke objects as such lack many ecologically relevant causal powers, and therefore have less bearing on our welfare than k-objects. It follows that adding them to our ontology runs against principles of simplicity while bringing little explanatory benefit. It is plausible, then, that the objects discerned and tracked by the visual system are k-objects rather than Spelke objects. But it seems unlikely that visual representations represent k-objects *as* members of kinds, for the membership criteria for kinds tend to privilege deep explanatory characteristics rather than the more superficial characteristics that are accessible to perceptual systems. A natural conclusion, then, is that perceptual representations represent k-objects, but only *as* bearers of congeries of low-level properties that are strongly indicative of kind membership.

X. Conclusion

While particularism fails completely to account for hallucinatory experience, existentialism can fully explain the phenomenology of hallucinations, their causal powers, and their role in giving reasons for thoughts and deeds. Given that these features of hallucinations are shared by veridical states, it is best to believe that both hallucinations and their veridical brethren have a quantificational semantics. It follows that awareness of particular objects cannot be fully explained in terms of the semantic content of perceptual experience. Instead we must explain it in terms of externalist conditions having to do with relations of satisfaction and causation between perceptual experiences and objects in the environment. The chapter offers a suggestion as to how that explanation might go. The other main conclusion of the chapter is that the ontology of perceptual experience—the class of objects over which its quantifiers range—consists of ordinary physical objects, partitioned into natural kinds. It would be an unnecessary complication of our ontology to suppose that perceptual objects are sui generis, in the way that the writings of perceptual psychologists sometimes suggest.

5
Perceptual Phenomenology

I. Introduction

One of the most important facts about perception is that it has a *phenomenological* dimension—in other words, a dimension that consists of *phenomenal properties* or *qualia*. As we observed in Chapter 1, perceptual qualia are qualities like tasting sweet, smelling of cinnamon, feeling smooth, sounding high pitched, sounding loud, looking brown, looking spherical, looking small, and looking distant. They are the ways that objects perceptually appear or seem to conscious observers (Kim 2006, 225). As such, they are plausibly regarded as properties of external objects (it is this orange that tastes sweet, and that tennis ball that looks spherical), but they are different from the objective physical properties of such objects. A wall can look dark brown even if it is tan, due to the dimness of the illumination; a planet can look small even if it is in fact large, due to its distance from the observer; and the rim of a cup can look elliptical even if it is circular, due to the angle from which the observer is viewing it. The point of the present chapter is to analyze the qualitative dimension of perceptual experience. I will focus on visual experience, but I believe the conclusions I will reach apply to the phenomenal properties of other forms of perceptual experience.

Given that I have just now identified visual qualia with visual appearances, it might seem that I have already proposed a theory of visual phenomenology, back in the chapters on appearance and reality; for those earlier chapters argue that ways of appearing can be identified with certain relational, viewpoint-dependent properties. In fact, however, the earlier proposal leaves some extremely important questions about perceptual phenomenology open. To see the kind of thing I have in mind, consider the fact that when we are aware of how things appear to us, it does not seem that we are aware of relational, viewpoint-dependent properties of the given sort. When, for example, an object looks small to us, we have no inkling that we are aware of a relational fact that derives from applying functions to the visual angle that the object subtends with respect to the lens of the eye. But it can seem that if looking small really is a relational property of this sort—if that is all that there is to it—we should have at least some inkling of this fact when something looks small to us. Or, more precisely, we should have some inkling of the fact unless it is true that, instead of being directly aware of the apparent size of the object itself, we are only aware of an *appearance* of the apparent size. It might be that an appearance of the apparent size fails to reveal all of the true nature of

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0005

the appearance itself, just as the appearance of water fails to reveal the underlying complexity of its chemical structure. But this seems absurd! How could there be an appearance of an appearance?

Here, then, is a difficulty about perceptual phenomenology that is not solved, or even addressed, by the theory of perceptual appearances that is put forward in earlier chapters. Another difficulty is posed by the fact that perceptual qualia are not the only qualia there are. In addition, there are sensory qualia (qualia associated with bodily sensations, such as pain), emotional qualia (such as the way it feels to be angry), and imaginative qualia (such as the impressions of color one has when one is imagining a national flag). Arguably there are also cognitive qualia—qualia associated with conscious thoughts. Qualia of these other types are all different from perceptual qualia in various ways, and are also different from each other, but they are like perceptual qualia, and like each other, in two fundamental respects: first, we have first person access to them, and, second, insofar as we are occupying the vantage point of folk psychology, it *seems* to us that qualia cannot be fully grasped or understood in any way other than the first person perspective. To see that the examples of qualia I have mentioned fit this characterization, consider the appearance of red—the way red things look to normal observers. This quale can be grasped from the first person perspective, and when we so grasp it, we become strongly inclined to reject the idea that it might be grasped as well from a third person perspective. Thus, it seems absurd that it might be fully grasped by someone who was color blind. Similarly, it seems to us that someone who had never experienced pain could not fully grasp its intrinsic nature. It would be no use to try to enlighten the person by describing the behavior that pain causes, or by explaining that pain is caused by damage to the body, or by providing information about the neural state that is correlated with pain. To grasp pain fully one must experience it oneself. Or so it seems.

I have said that it *seems* to us that qualia cannot be apprehended from a third person perspective. I believe this succeeds in delimiting the intended class, thereby providing a useful basis for generalizing about its members, but I don't think that the fact that we have this impression about qualia is a trustworthy guide to their essential nature. I say this because there are plausible arguments for reducing qualia to physical properties of various kinds, and, of course, physical properties can always be grasped from the third person perspective, if they can be grasped at all. That is, there are plausible arguments that qualia are properties that can be grasped both from the first person perspective and also from a third person perspective (e.g., the perspective of neuroscience). We don't want to rule out such views by fiat, as we would be doing if we were to specify the class of qualia by saying that they are properties that can *in fact* only be known fully from the first person perspective.

The impression that qualia can only be apprehended from the first person perspective is fostered by folk psychology. (How this happens is something that we

will consider soon.) If we were to view the impression as veridical we would in effect be making folk psychology the arbiter of questions in the philosophy of mind. Far better to weigh pronouncements of folk psychology against those coming from other vantage points on mental states and processes, especially the vantage points of psychology, neuroscience, and scientifically informed philosophy. (For a quite different treatment of the impression, see Nagel 1979.)

How then should we pick qualia out when we wish to generalize across all of them? I suggest that we won't be in a position to answer this question in any final, non-provisional way until, guided by science and philosophical reflection, we have moved beyond folk psychology to a more adequate view of the functional and intrinsic properties of paradigmatic qualia like the ones associated with perception and interoception. Eventually, we will be able to see, I think, that while qualia share certain properties, what principally unifies them is that they can all be apprehended (or imagined) using certain types of internal representational scheme. In other words, the unity of qualia is to a large extent second hand, deriving from the unity that characterizes the forms of awareness that are directed on them. The representational schemes that support these forms of awareness have formal and functional commonalities that will gradually become apparent as we proceed, beginning in the present chapter and continuing through the next three. (If we think of awareness of phenomenology as a natural kind, it should not surprise us that we have to begin with a suboptimal characterization. It is generally true that in arriving at an adequate specification of a kind, we begin with a commonsense or folk description that is eventually replaced, giving way to a deeper and more explanatory characterization.)

In Section II I will review the challenges that perceptual phenomenology poses for attempts to develop a unified, integrated view of the natural world. Section III will set out some of the main claims of the theory of perceptual phenomenology that I wish to propose. These claims will all be concerned with the nature of *awareness of* phenomenology. Roughly, the overarching idea is that the main metaphysical problems involving perceptual phenomenology can be solved by adopting a representationalist theory of phenomenal awareness. Section IV will present arguments on behalf of such theories, and Section V will address a popular objection to them. Appendix I and Appendix II will extend the account of awareness of perceptual phenomenology to awareness of cognitive qualia and awareness of affective qualia, respectively.

As this outline indicates, the present chapter will be mainly concerned with questions about *awareness* of perceptual qualia. It won't have much to say about qualia themselves. But metaphysical questions about the nature of qualia are addressed in the appendices and in other chapters. The chapters on appearance and reality have already recommended a theory of the nature of perceptual qualia, and the next chapter will present a theory of pain qualia that will generalize to other somatic qualia.

II. Dualism

In this section I will review several plausible arguments that qualia are metaphysically sui generis, in the sense that they are not reducible to or grounded in physical phenomena of any kind. If these arguments are sound, then nature is bifurcated into components or aspects that may be connected by laws of nature but are metaphysically independent: each could in principle have come into existence without the other, and each could in principle disappear without affecting the other. Thus, for example, the arguments imply that insofar as the Big Bang was a purely physical event, it cannot explain all aspects of the natural world. Specifically, it cannot explain the existence of qualia or the laws linking qualitative states to neural states. Many people find it uncomfortable to think of nature as having this sort of irreducible duality, though others welcome the idea, seeing in it a hint that it is possible for a conscious being to survive the death and disintegration of the physical body.

The knowledge argument: Consider someone—call her "Mary"—who has never seen a red object or even imagined one, but who is otherwise immensely well informed about color and color perception. In fact, Mary knows every physical fact about color, including all facts about the surface reflectances of objects, all facts about the wavelengths of light, all facts about the retina's ability to pick up information about wavelengths, and all facts about the processing of such information in the brain. But since she has never seen a red object, she has no grasp of how red objects appear to a normal human observer. Suppose now that Mary sees a red object for the first time. She will then grasp something that she has never grasped before—how red things look. But if she now knows how red things look for the first time, and she formerly had a deep knowledge of all of the physical properties that are instantiated when someone sees a red object, then it cannot be true that the way red things look is reducible to a physical property or a congeries of physical properties (Broad 1925, Jackson 1986).

The Cartesian modal argument: The imagination provides our principal and most fundamental test for possibility: if we want to know whether an object or state of affairs is possible, we see if we can coherently imagine it. Moreover, we regard it as a highly reliable test. Now it is possible to imagine perceptual experiences with phenomenal properties existing without being accompanied by physical phenomena of any kind. Putting these ideas together, we obtain the result that it is possible for perceptual phenomenology to exist without being accompanied by physical phenomena. But if it's possible for X to exist without being accompanied by Y, then X cannot be identical with Y, and more generally, it cannot depend metaphysically on Y (Descartes 2017, Kripke 1981, Chalmers 1997).

The unbridgeable gulf argument: In paragraph 412 of his *Philosophical Investigations*, Wittgenstein (1953) writes as follows:

> The feeling of an unbridgeable gulf between consciousness and brain-process. . . . This idea of a difference in kind is accompanied by slight giddiness. . . . When does this feeling occur in the present case? It is when I, for example, turn my attention in a particular way on to my own consciousness, and, astonished, say to myself: THIS is supposed to be produced by a process in the brain!—as it were clutching my forehead.

Wittgenstein goes on to disparage the ideas of attention and consciousness that he believes to provide a foothold for this sort of vertiginous aperçu, but the problem the passage describes is much more compelling than his attempt to dissolve it. When we are aware of a brain state by external observation of some kind, we are aware of entities and properties that seem to be complete in themselves, lacking any trace of the appearance of red or the appearance of circularity. And when we are introspectively aware of phenomenal properties, our awareness reveals no trace of the neural entities and properties in question. More specifically, it seems that we can fully grasp an appearance of red without having any indication whatsoever that the appearance might be intimately related to neurons, action potentials, and the release and binding of neurotransmitters. In short, perceptual appearances seem to be altogether different from the brain processes that accompany them. Related arguments show that appearances also seem altogether different from all of the other physical phenomena with which they are associated, such as the Thouless properties we considered earlier. How then could phenomenal properties be in any way reducible to physical properties (Levine 2001)?

If these arguments are sound, then perceptual qualia mandate a dualist perspective on reality.

As stated here, the arguments are concerned exclusively with perceptual qualia, but they can be generalized so as to support dualist claims about qualia of other kinds, such as sensory qualia like pain, and emotional qualia like the way it feels to be afraid. For example, in its generalized form, the knowledge argument yields the conclusion that someone who has never felt pain is unable to grasp its qualitative nature, no matter how much the person may know about the neural phenomena that accompany pain. This conclusion stands or falls with the corresponding conclusion about Mary's grasp of the reddishness of perceptual experiences.

We should note that all of the foregoing arguments presuppose that there is no appearance/reality distinction defined over qualia. More specifically, they presuppose that perceptual qualia have no "dark side" or hidden microstructure. For consider. If the way red things look were not fully accessible to awareness, then when Mary comes to appreciate how red things look for the first time, it could be that there is something more to the property she is aware of than her awareness of it reveals, and that hidden dimension could have the structure of a Thouless property. Equally, when a Cartesian philosopher tries to imagine a situation that contains an instance of phenomenal red but no neural phenomena, it could be that

the imagined instance of phenomenal red has a neural aspect that the philosopher's imagination has not captured. And when a philosopher moved by the passage from Wittgenstein introspectively appreciates that something looks red, but finds no hint or trace of neural phenomena in this appearance of red, it could be that the philosopher has failed to grasp the deeper complexity of the appearance, due to the simplicity of the appearance of the appearance. The arguments for dualism seem quite safe because it seems to us that in the case of perceptual qualia, what you "see" is what you get. But if this background assumption were false, it would be a wholly different story.

III. Phenomenal Representationalism

In this section I will describe *phenomenal representationalism*, a theory of awareness of phenomenal properties that incorporates an appearance/reality distinction. It consists of seven theses.

1. All awareness is representational (i.e., all awareness involves one or more representations of the object of awareness).

I will explain the grounds for this assumption in the following section. For now, I will just say that it has turned out to be an immensely fruitful assumption in cognitive science.

2. Representations can distort the items that are represented, and can also fail to register their essential natures.

Examples of distortion include a watch that is running too fast and a photoshopped picture of a political rival, which makes the rival look demonic. In the realm of natural representation, body dysmorphic disorder distorts the flaws in one's appearance, making them seem much more prominent and apparent to others than they are, and asomatognosia distorts one's image of the extent of one's body, making it seem that large parts of it are missing.

For an illustration of how representations can fail to capture the essential natures of the items they represent, consider the case of line drawings. To be more specific, consider a line drawing of a specific house—say, a drawing of the front of the house at 25 Elm Street (Figure 5.1).

This drawing represents the door and windows on the first and second floor of the house, and also the door, roof, and chimney. But it doesn't represent the materials from which the house is constructed. For all the drawing reveals, the house could have been made of bricks, or cinder blocks, or field stones, or adobe, or a number of kinds of wood. This is a limitation of line drawings as a system of representation: they necessarily abstract away from the material compositions of things. But this means that the line drawing of the house at 25 Elm Street fails to do full justice to its essential nature, for the house could not have been made of materials other than the ones that actually compose it (Kripke 1981). Now

Figure 5.1. Drawing of the house at 25 Elm Street.

consider a line drawing of the floor of the living room of the house. Suppose it simply represents the floor as rectangular, without providing any information about the mereological structure of the floor. Thus, it leaves it open whether the floor is solid or is instead made of planks of some kind. As in the case of the first drawing, this one also fails to do full justice to an aspect of the essential nature of the item it represents, for whether the floor is solid or divided into long rectangular strips of some material is a constitutive fact about it. Of course, a different line drawing of the floor might have told a more illuminating story. Line drawings are *capable* of registering mereological structure. But the one we are considering does not. In general, a line drawing can only represent parts of an item by having parts of its own that correspond to those parts.

Here, then, is an example of a representational system that can fail to capture two kinds of essential property—material composition and mereological structure. For examples involving natural representations, consider the fact that human visual representations cannot register the molecular structure of water and the fact that haptic representations often fail to register the material composition of surfaces. ("It's very smooth, but is it plastic or porcelain or highly polished wood?" When our eyes are shut, we are generally unable to answer such questions.)

3. Because representations can be erroneous and fail to capture essential natures, it is always possible, on a representational account of awareness, to draw an appearance/reality distinction with respect to the object of awareness—there is the object as it is represented (appearance), and the object as it is in itself (reality).

The type of appearance reality/distinction that I'm drawing here was emphasized by Kant, so I will sometimes describe it as the *Kantian* distinction in the sequel.

4. Since all awareness is representational, awareness of qualia is representational. Hence, awareness of qualia can distort qualia and also fail to register their essential natures.

Folk psychology fails to recognize the representational nature of awareness of qualia, and as a result, it fails to recognize that awareness of qualia is governed by a Kantian appearance/reality distinction. Thus, for example, folk psychology will not allow us to distinguish between how pain seems to us and how it is in itself. As a special case of this, since the sensory character of pain and the hurting-ness of pain are indissolubly fused in our everyday experience of pain, folk psychology will not allow us to entertain the possibility that these characteristics are distinct in reality, and are in fact dissociable. But we know from the testimony of asymbolia patients that this is in fact the case. How pain seems to us in everyday experience is not the way it is in reality. (Asymbolia and its implications are discussed at length in Chapter 6.)

The blind spots we've just been considering are among the most serious failings of folk psychology. In my view, they are largely responsible for the illusion that awareness involves an acquaintance relation. Also, as I will explain in a moment, they are responsible for the mind–body problem.

5. Since awareness of qualia is representational, there is an easy explanation of why perceptual awareness of Thouless properties fails to do justice to their relational nature. The explanation falls out of the general principle that perceptual representations do not as a rule capture the full character of the items they represent. As we saw, this is true of visual representations of water, and of haptic representations of smooth surfaces. It is also a familiar fact that perceptual representations can fail to do justice to the relational character of properties. Consider, for example, our perception of the weight of an object. It fails altogether to register the fact that weight isn't a property of a single object, but rather consists in a gravitational relationship between objects and the Earth.

6. The Kantian appearance/reality distinction also calls the arguments for property dualism into question. As we saw, they all presuppose that it is impossible to draw an appearance/reality distinction with respect to awareness of phenomenal properties. We now know, however, that this presupposition is false. It could well be, for example, that awareness of qualitative states involves representations that fail to reflect the material composition of the states, and also their fine-grained mereological structure. (See thesis 2.) Accordingly, it is possible to respond to the arguments as follows.

The knowledge argument: When Mary perceives a red object for the first time, her visual system tokens a representation that has never been activated before. This allows us to explain why it seems to her that she is aware of something new without supposing that this impression is veridical. It seems to her that she is

aware of something new because she is representing a familiar item, the appearance of red, in a dramatically new way. (Formerly she could only use conceptual representations of the appearance of red, perhaps supplemented by black and white images of various kinds.) In other words, it seems to her that she is aware of something new because the object of awareness is presenting a different appearance, in the Kantian sense, than it had in previous encounters. The appearance of red is appearing differently to her.

The Cartesian modal argument: As we saw, this argument depends on the claim that we can imagine perceptual phenomenology that is not accompanied by any physical phenomena. But how do we know that we can imagine this? Well, it certainly *seems* that we can imagine something's looking red without thereby imagining a Thouless property. But this is only how things *seem*. Maybe what we're imagining *really is* a Thouless property. At this point in the dialectic, a Cartesian will say: "But in the case of the appearance of red, there is no difference between how things seem and how they really are. There is no appearance/reality distinction defined over qualitative states. So it *couldn't* be the case that when you are imagining that an object presents an appearance of red, you are imagining something that might really be a Thouless property." Historically this argument has seemed unanswerable, but representationalism can challenge it, since representationalism can draw an appearance/reality distinction with respect to awareness of Thouless properties.

The unbridgeable gulf argument: To evaluate this argument, let's for a moment turn our attention from perceptual qualia and Thouless properties to the sensory quale *pain* and the neural processes that accompany it. This will show that phenomenal representationalism, which has up to now been presented as a theory of perceptual qualia, can be generalized so as to undercut dualistic arguments that are concerned with qualia of other types.

Why does there seem to be an unbridgeable gulf between nociceptive neural activity in regions of the body and pain? I suggest it is because the representation that is deployed when we are introspectively aware of pain is different from all of the representations we use in perceiving or thinking about phenomena that are allegedly distinct from pain. The representation that subserves introspective awareness of pain is a component of a system of representation that we use *only* when we are tracking our own bodily states "from inside." Now it is a general rule that if we are aware of a phenomenon X by virtue of a representation that is different than the representation that supports our awareness of Y, it will remain an epistemic possibility for us that X is distinct from Y until we find grounds for thinking that they are identical, such as evidence that X and Y have the same spatio-temporal location. In the present case, the fact that our introspective awareness of pain is supported by a special representation creates the epistemic possibility that pain is distinct from nociceptive neural activity.

Now it often happens that we are aware of a single phenomenon via different representations. For example, we can be aware of the roughness of a surface via both visual and tactual representations. In most of these cases, we are able to go beyond the epistemic possibility of difference to conclude that we are aware of a single item in two different ways. But this is not what happens in the case of introspective awareness of pain. In that case, we are tempted to move from the epistemic possibility of difference to the metaphysical claim that pain is distinct from all other phenomena. And we go on to infer that pain can only be apprehended in one way. Why do we proceed differently in this case than in others? I think it is clear what the answer is. We conclude that pain is distinct from everything else because folk psychology in effect denies that our awareness of pain is supported by a representation. According to folk psychology, there is no difference between appearance and reality in the case of pain. Rather, introspective awareness of pain is a matter of direct acquaintance. It follows that what is given in introspection is all there is to the reality of pain. More particularly, it follows that pain could not have the properties of an independently apprehended item without our appreciating that it has those properties.

To summarize: Folk psychology denies that introspective awareness of pain is representational, and, in doing so, it forces us to conclude that there is an unbridgeable gulf between pain and everything else. And this leads naturally to the view that we can only be aware of pain in one way. Fortunately, there is an easy way of avoiding these conclusions—we need only reject the theory of awareness of qualia that is implicit in folk psychology and replace it with the representationalist picture that is suggested by cognitive science.

7. It is possible to explain (or explain away) the special properties of perceptual qualia that seem to make them non-physical by appealing to special properties of our ways of representing them.

Consider, for example, the property *looking yellow*. This property seems absolutely simple to us, but there is nothing simple about the physical properties that are associated with looking yellow. Yellow light consists of a complex vector of wavelengths; the processing of yellow light by cone cells in the retina is hugely complicated; and the same is true of the neural activity in the upper visual system that supports perceptual consciousness of yellow. So we are led to ask, how could *looking yellow*, which seems absolutely simple, be a physical phenomenon?

Representationalism can answer that question as follows: Suppose that visual representation has an iconic dimension. Now we know it to be a law of iconic representation that complexity in the represented object is represented by complexity in the representation itself. Thus, for example, if a line drawing of a house lacks rectangular components corresponding to bricks, then it does not represent the house as composed of bricks or brick-like elements. Applying this law of iconic representation to the case of visual representation, if an iconic visual representation fails to reflect the complexity of a physical property, it doesn't represent

that complexity. It fails to articulate it, and it therefore fails to represent it. This doesn't mean that it represents the property as simple. Failing to represent as complex is not the same thing as representing as simple. But from the perspective of a subject whose cognitive command of a physical property derives entirely from a visual representation, R, R's failing to represent the complexity of the property will come to the same thing as attributing simplicity to the entity. Because the subject's cognitive command of the entity derives entirely from the representation, the subject will find it natural to attribute simplicity to the entity in thought even if the perceptual representation does not.

In other words, it is plausible that *looking yellow* seems simple to us because, at the most fundamental level, our awareness of it derives from a simple representation. Since the representation is in effect an atomic component of the system of representation with which it is associated, it will provide the subject who is using it with no ground for attributing complexity to yellow. Accordingly, it will be natural for the subject to form the impression that the represented property is simple. (A representation belonging to a system of representation Σ is *atomic* relative to Σ if it cannot be factored into parts that are themselves representations belonging to Σ. A state or process that counts as an atomic representation relative to Σ may still have parts in a physical or spatial sense. It's just that they can't have the status of representations (relative to Σ).)

Notice that this account of apparent simplicity makes a crucial appeal to a syntactic property of our perceptual representational of yellow—that it is a representational atom. It follows that on phenomenal representationalism, the theory of phenomenal awareness I am recommending, it isn't possible to give a full explanation of our *impressions* of qualia in terms of qualia themselves. Rather, we have to appeal to features of the representations that stand for qualia. Another way to put this is to say that while phenomenal representationalism maintains that qualia are objects of awareness, it also claims that certain aspects of the way qualia appear to us have to be explained in terms of modes of awareness.

I emphasize that, like other representationalists, my view is that qualia figure in the representational contents of states of awareness. They are objects of awareness. As I see it, however, in addition to being aware of qualia, we also have impressions of what qualia are like in themselves (i.e., beliefs about the natures of qualia), such as the impression that the quale *looking yellow* is simple. These impressions are fostered by the intrinsic, syntactic characteristics of perceptual representations. Many of them are correct, but others are misrepresentations. (For further discussion of the relationship between representational format and representational content, see Lande 2018.)

This completes my exposition of phenomenal representationalism. To summarize, its underlying idea is that the metaphysical problems that are generally regarded as posed by phenomenal properties are actually due to two flaws of folk psychology—its failure to recognize that awareness of phenomenal properties has

a representational structure, and its attendant blindness to the fact that such awareness is governed by a Kantian appearance/reality distinction. The key to solving the mind–body problem is an adequate theory of phenomenal awareness.

This emphasis on awareness of phenomenology is one of the main differences separating my position from the position of philosophers like Ned Block, who holds that qualia are properties of perceptual experiences themselves rather than components of the representational contents of experiences. (See, e.g., Block 1999.) As far as I know, Block has never proposed a positive theory of awareness of phenomenology. This emphasis on awareness also distinguishes my position from that of Tyler Burge, who seems to hold that qualia play a role in the perception of external objects that is similar to the role that senses play in Frege's philosophy of language (Burge 2010). Like Fregean senses, they are modes of presentation—ways in which external objects are presented to minds. More concretely, insofar as Burge recognizes perceptual qualia at all, he seems to view them as modes or ways of being perceptually aware of external objects. It follows that they are external to the contents of the representations that constitute perceptual awareness. How then do we become aware of them? As with Block, I know of no place in Burge's writings in which he addresses this question. Moreover, I am pretty sure that if he were to address it, he would wind up in the snares associated with the "adverbial theory" of phenomenal awareness. I will describe the theory and the attendant snares in the next section, and say more about this matter in Chapter 7.

IV. Reasons for Representationalist Accounts of Awareness

The first thesis of phenomenal representationalism is that all awareness constitutively involves representations. I will defend this thesis in the present section. The defense will have two parts. I will begin by pointing out that there is a substantial empirical case for the thesis. After that, I will construct an argument by cases designed to show that representationalist accounts are superior to their competitors.

As noted earlier, representationalist accounts of awareness have proved immensely fruitful in cognitive science. References to representational states are pervasive in theories of high-level mental states and processes: all such theories are concerned with memories, beliefs, plans, decisions, analyses of problems, emotions, and so on, and it is generally true that such theories imply that the states and processes with which they are concerned have a representational dimension. For example, a memory is treated as a representation of something that happened in the past, and anger is seen as involving a representation of something that is an obstacle or poses a threat. Theories of perceptual processes are also deeply and pervasively committed to representationalism. Today this is

true even of Gibsonian accounts of perception. In earlier times Gibsonians were deeply skeptical about the need for internal representations, but that attitude was generally due to a misunderstanding: they assumed that representations had to be internal models of external phenomena, that representationalist accounts of perception had to be committed to some sort of awareness of such internal models, and that, according to representationalism, awareness of external phenomena is indirect, being achieved only by inference from the properties of the models (Gibson 1979, Warren 2005). In fact, however, instead of viewing representations as objects of awareness, contemporary representationalists think of them as constitutive of awareness. Perceptual representations are not signs that we observe and interpret, but rather components of perceptual awareness itself—components that make the external directedness of awareness possible by having contents in which external phenomena are involved. Today many Gibsonians acknowledge that perceptual processing must consist of computations defined over information that is registered by internal states, where these states really come to the same thing as representations. The difference currently separating these Gibsonians from other perception scientists is just that the former tend to place more emphasis on the amount and quality of information about distal phenomena that is contained in proximal stimuli, maintaining that the proximal information is sufficiently rich that the internal equipment needed to transform it into representations of distal phenomena can be comparatively modest (Warren 2021). The very general hypothesis that perceptual awareness involves representations, and that perceptual processing consists in computations on representations, is, as far as I can tell, shared by the overwhelming majority of perception scientists. And it is absolutely thrilling to contemplate the immensely rich portrait of perception that this hypothesis has enabled them to construct! (Textbooks provide the most efficient means of access to this portrait. See, e.g., Palmer 1999, Wolfe et al. 2006, Frisby and Stone 2010, Yantis 2014.)

I will now give an argument by cases for representationalist theories of awareness by criticizing alternative accounts. There are three such alternatives. One maintains that experiential awareness is a matter of *direct acquaintance* with facts involving objects and certain special properties, where acquaintance is thought to be a primitive cognitive relation that obtains between observers and certain facts. On this view, awareness is simple and unstructured, and, in particular, it does not have representations as constituents. It occupies the ground floor with respect to explanation and analysis of perceptual awareness. A second alternative to representationalist theories is *adverbialism* (Ducasse 1952). Roughly speaking, adverbialism claims that when a subject X is aware of an entity Y as having a quale Q, what is going on is that X is aware of Y in a certain *way*. To be a bit more specific, according to adverbialism, X is aware of Y Q-ishly. Thus, while adverbialism allows that qualia exist, it maintains that they exist only as *forms of perceptual awareness*—as *ways* of being perceptually aware of non-qualitative phenomena.

That is to say, adverbialism maintains that qualia are adverbial qualifications of an underlying generic relation of awareness, a relation that agents bear to external objects. The third alternative to representational theories of experiential awareness consists of what might be called *doxastic theories*. According to theories of this stripe, qualitative awareness is propositional and doxastic in character, and therefore necessarily involves some sort of conceptualization. To be aware of a quale is to token a representation of a property, but the representation in question is fully conceptual in character. All knowledge of qualia takes the form of judgments.

Acquaintance theories. There is a notion of acquaintance that figures prominently in folk psychological explanations and in everyday conversations that is free from the assumptions noted above. Using this notion, we can say that Kate is acquainted with all of the restaurants in her neighborhood. Equally, we can explain why Kate made a rude gesture to a politician by saying that she is well acquainted with his record of cowardice and duplicity. In such uses, we are deploying a notion of acquaintance that is fully compatible with—and is in fact best explicated by—representationalist accounts of awareness. The notion of acquaintance that figures prominently in contemporary philosophy is quite different. It was originally fashioned by Bertrand Russell, who was moved principally by epistemological considerations (Russell, 1911, 1912, 1914). Russell wanted a foundation for knowledge, and believed that to attain this end, it is necessary to posit a mode of awareness that is free from all forms of interpretation and is completely independent of inference. For interpretation and inference introduce the possibility of error. It follows that Russellian acquaintance is free from conceptualization and all other forms of representation. Moreover, it must be a simple relation, for if it had an internal structure, that structure could impair our access to external objects, much as the muntins that separate the panes of glass in a window can interfere with our view of what is outside. One might respond to this reconstruction of the ordinary notion of acquaintance by saying that knowledge doesn't need such a secure foundation: we can have knowledge of a topic even though our grasp of it is partial or grounded in probability. This is, at all events, the response that seems right to me. But Russell seems to have thought that fallibilistic theories of knowledge lead inevitably to skepticism, and that his bullet-proof relation of acquaintance represents our only hope of blocking skeptical views.

A number of contemporary philosophers have embraced Russellian acquaintance. This allegiance is no doubt partly due to Russellian epistemological concerns, but it may also be due to an appreciation of the fact that folk psychology implies that awareness of qualia is bullet-proof, allowing no possibility of error or incompleteness. That is, although the ordinary notion of acquaintance seems to be much less freighted with requirements motivated by epistemological concerns than Russell's notion, it seems that folk psychology also contains a concept of

awareness that corresponds more closely to what Russell had in mind. To be sure, this second concept is restricted in its application to awareness of qualia, but if one thought, as some contemporary Russellians appear to do, that qualia include objective properties of external entities, such as objective colors and objective shapes, one might also think that perceptual awareness of the environment meets the demands of Russellian acquaintance. At all events, whatever the rationale, and whatever the etiology, the idea that perceptual awareness conforms to more or less the requirements of Russellian acquaintance is well entrenched in contemporary philosophy.

As I see it, however, there are strong reasons for preferring representationalist accounts of awareness to accounts that are based on acquaintance.

One is that representationalist theories provide a basis for answering certain key questions about the nature of awareness itself. At any given time one is aware of certain objects and characteristics and not of others. Why is one aware of the items that one is in fact aware of? Why isn't one aware of the others instead? Further, why is the scope of awareness broad at some times and not at others? Why does the resolution of awareness change when it does? What is the difference between attentive awareness and more casual forms of awareness? We are not yet in a position to give detailed answers to questions of this sort by appealing to properties of the system of representation that supports experiential awareness, but elementary reflection shows that it is possible in principle to answer them in this way. Thus, for example, it is clearly possible to explain differences in the scope of awareness by appealing to differences in the contents of representations: the more inclusive the content of the representations that are involved in awareness, the broader the scope of awareness. It is also possible to explain differences in resolution in terms of differences in the contents of representations. Thus, if experiences have representational contents, it's comparatively easy to explain why my current experience of my thumb reveals the whorls of a thumbprint. It reveals the whorls because it is constituted by a representation that represents the whorls. The acquaintance theory has trouble explaining facts of this sort. Consider, for example, what an acquaintance theory would say about experiential awareness of my thumb—that it has my thumb and perhaps myself as constituents, but that it is otherwise fundamental, lacking representational structure and internal structure of any other kind. How can this hypothesis explain the fact that I am aware of the whorls of the thumbprint, but not of the microphysical structure that is no less a property of my thumb? And how can it explain the fact that if my thumb was farther away, or I was more nearsighted, I would not be aware of the whorls? The answer is that it can't. Acquaintance theories attribute too little structure to perceptual awareness to be able to explain even its most basic features.

Acquaintance theorists sometimes try to evade this objection by saying something like, "Well, facts of acquaintance depend causally on processes in the brain, and in the present case, those brain processes are shaped by detailed information about the whorls on your thumb." In other words, advocates of the acquaintance

theory sometimes try to pass the buck to another level, maintaining that the facts in question have nothing to do with the nature of awareness itself. If we want to explain them, we must appeal to neural processes and make a claim like, "Well, there are primitive psychophysical causal laws linking neural processes to facts of acquaintance." But to say this is in effect to acknowledge that *the acquaintance theory has no explanatory power of its own. It can't explain things like focus, perspective, and resolution in terms of acquaintance because it doesn't attribute any positive properties to acquaintance.* The acquaintance theory contrasts sharply in this respect with representationalism, for representationalism can explain aspects of awareness like focus, perspective, and resolution quite nicely without appealing to irreducible facts and primitive laws.

In view of these considerations, it is clear that representationalist theories of perceptual awareness are overwhelmingly superior to those based on acquaintance in point of explanatory power. This is a good reason for preferring them. A second reason is this: if perceptual awareness lacks all internal structure, then there can be no ground for the fact that perceptual awareness has various important causal powers involving internal phenomena—that is, no ground for the fact that perceptual awareness contributes causally to many different mental processes, including the fixation of beliefs, the establishment of memories, the formation of plans, and the triggering of emotions. Thus, according to acquaintance theories, facts of perceptual awareness have just three constituents—a subject of awareness, an object of awareness, and a relation linking the two. The only distinguishing feature of such a fact on the subject's side is the subject herself. Hence, once the subject has been fixed, facts of awareness have no differentiating characteristics other than the objects of awareness. There are no internal states of the subject, and, in particular, no representations, that contribute to differentiating the facts. But we know, for example, that perceptual awareness of a green object has very different causal implications for the internal economy of a subject than perceptual awareness of a red object. If a subject is perceptually aware of a green object, the subject will very likely form the belief that the object is green, as opposed to the belief that the object is red. Now there must be an internal trigger for a causal process that results in a new belief. Given that facts of awareness can cause beliefs, it follows from this that facts of awareness must involve internal triggers of some sort. Representations can play this role, but acquaintance theories make no allowance for triggers. States of acquaintance have no place in the functional architecture of the mind.

Some acquaintance theorists have tried to meet this objection by modifying the basic Russellian picture of acquaintance. On their view, acquaintance is a three-place relation linking a subject, an external object or state of affairs, *and* an internal state of the subject (Martin 1998, Beck 2018). The properties of the internal state are supposed to capture the way that the external entity appears to the subject. In effect, they are supposed to be perceptual qualia, perhaps reducible to

neurocomputational properties of some kind. Now the fact that acquaintance has this third term makes it possible for there to be different facts of acquaintance even when the subject and the object of acquaintance are held constant. Moreover, since the third term is internal, differences of this kind make it possible for facts of acquaintance with the same subject and the same object to make different contributions to internal processing. Allegedly, this answers the foregoing objection.

But the theory is untenable. Reflection shows that it is equivalent to the adverbial theory of perception, which I will be discussing soon. Accordingly, it fails for the same reason as that better-known theory. Another issue is that it seems not to secure the epistemological benefits that acquaintance is believed by naïve realists to confer. It suggests that internal states color our awareness of external phenomena. That doesn't fit together comfortably with the Russellian idea that perception affords an unvarnished grasp of objective properties.

A third problem with acquaintance theories is that the very considerable success of contemporary vision science speaks loudly against the existence of a primitive, unanalyzable relation of acquaintance. The achievements of vision science are largely due to the assumption that perceptual awareness isn't primitive or unanalyzable, but is on the contrary susceptible of deep analysis and explanation.

It might be replied that acquaintance theories of perception are not in competition with vision science—rather, vision science aims only to explain the *causal basis* for acquaintance and other higher-level perceptual phenomena. (See, e.g., McDowell 2013, which proposes a "two levels" view of perception, where the bottom level consists of causal infrastructure and the upper level is occupied by agent-level perceptual phenomena.)

Although this alternative view has a certain currency, it does not survive an examination of the actual ambitions and results of vision science. As any such examination will reveal, vision scientists are concerned with very high-level questions. They aim to analyze seeing itself, not merely a causal substratum of seeing. Thus, as the literature shows, vision scientists are very interested in questions about conscious visual experiences, including questions about such experiential topics as the relationship between appearance and reality, the tracking of objects across time, the nature of perceptual space, the phenomenology of colors, and visual attention to locations, objects, and features. Moreover, they do not seek merely to give causal explanations of such experiential phenomena, but to say what the phenomena consist in. For example, "How exactly do objects appear to subjects?" is a question of much current interest (see, e.g., Carrasco et al. 2004, Durgin and Li 2017), as is the question of what geometry best describes phenomenal space (see, e.g., Todd et al. 2001, Wagner 2006). Of course, vision science is deeply concerned with lower-level, subpersonal states and processes as well, but it would be naïve to think that these lower-level phenomena are its only concern.

Another reason for preferring representationalist accounts of awareness is that the claim that perceptual awareness involves acquaintance appears to be

unverifiable. After all, acquaintance is defined almost entirely by negative properties. How could there be behavioral or imaging tests for something with virtually no positive properties?

I once heard an acquaintance theorist say that we know about acquaintance by being introspectively acquainted with it. But how do you know that you've succeeded in introspecting it, given that it has virtually no positive properties? Note that by definition it has no proprietary phenomenology.

When I consider these and other problems with the acquaintance theory, I am forced to conclude that it is quite the worst idea anyone has had since Thales declared that everything is water.

Adverbial theories. According to these theories, it will be remembered, qualia exist only as *forms of perceptual awareness*—as *ways* of being perceptually aware of non-qualitative phenomena. I see no merit in this suggestion. Adverbialism denies that there is such a thing as perceptual awareness of qualia. Instead, it maintains, qualia are ways of being perceptually aware of other things. But this claim fails to acknowledge a key fact. Surely we are aware of qualia. We are not simply ignorant of them, as we would be if there was no such thing as awareness of them. Clearly, for example, we know what it is for an object to look red. Hence, adverbialism must allow that there is non-perceptual awareness of qualia of some sort. Now this form of awareness will have to be explained in some way. But it can't be explained adverbially, on pain of regress. An explanation will have to appeal to acquaintance, or to non-conceptual representations, or to judgments. Thus, in the end, either adverbialism is unable to account for our knowledge of qualia, or it collapses into one of the other theories of qualitative awareness we have distinguished. I will say more about the problems of adverbialism in Chapter 7. (If I understand it properly, the proposal in French and Phillips 2020 is also vulnerable to this objection, even though strictly speaking the proposal is only a cousin of adverbialism, not a version of it.)

Doxastic theories. These theories allow that perceptual awareness constitutively involves representations, but they differ from the view I am defending under the label "representationalism" in that they claim that the representations in question are conceptual through and through. They maintain that to be aware of a phenomenal property is to make a conceptually informed judgment of a certain sort.

Reflection shows that doxastic awareness lacks a number of features that are possessed by perceptual awareness. These features include:

(a) Perceptual awareness provides us with access to highly determinate forms of properties. (When an object looks blue to me, it generally looks a highly determinate shade of blue.)

(b) It provides information about complex patterns that cannot easily be captured by linguistic or conceptual descriptions. (Try describing a Jackson Pollock canvas in concepts or words.)

(c) It enables us to make extremely fine discriminations among the members of domains of properties, thereby giving us a sense of the density of the orderings that obtain in the domains.

(d) It is associated with a variety of attention mechanisms, including mechanisms that enable us to adjust certain aspects of experiential awareness, such as resolution and figure/ground contrast.

Awareness that involves conceptualization and judgment lacks all of these features. It is not true, for example, that conceptual awareness generally provides us with access to highly determinate forms of properties. I can judge that something is blue without judging that it has a highly determinate shade of blue. Nor does conceptual awareness have property (b), as the Jackson Pollock example illustrates. It also lacks property (c). Anyone who has perceptual access to colors will appreciate that similarity orderings of shades are very finely graduated. We can, after all, discriminate a million different colors. On the other hand, a blind person, even one who has somehow managed to acquire an immense number of color concepts, will not automatically have a sense of how finely graduated such orderings are. Finally, conceptual awareness lacks property (d). To be sure, we do speak of attention in connection with conceptualization and judgment. For example, it is possible to attend in thought to a theorem in number theory. But this sort of attention is different in kind than the forms of attention that are involved in experiential awareness. Thus, in the case of doxastic attention, there is nothing that is strictly analogous to the changes in apparent contrast, apparent size, apparent motion, and apparent saturation of colors that accompany visual attention (Carrasco 2011). Qualitative perceptual awareness must be distinguished from doxastic awareness.

V. An Objection to Phenomenal Representationalism

In the present section I will address what has come to be known as the *Swampman objection* to representationalist accounts of perception and perceptual phenomenology. It comes to the fore when such representationalist accounts are combined with *selectionist teleosemantics*—the very attractive view that the representational capacities of our perceptual systems are to be explained in terms of biological functions.

As the reader will recall, there is a discussion of selectionist teleosemantics in Chapter 1. I will briefly recapitulate the relevant parts of that discussion here.

Roughly speaking, selectionist teleosemantics consists of two ideas. The first is the etiological theory of biological functions, according to which it is the function of an item X to do or cause Y if X was selected by evolutionary processes because X did or caused Y on a number of historical occasions, and doing or causing Y on those occasions contributed to the fitness of the agents who possessed X. The second idea is that perceptual representation can be explained in terms of the

concept of a biological function and the concept of information. More specifically, it is the idea that a perceptual state X can be said to represent an entity Y just in case it is the function of X to encode information about Y. There are of course many refinements and extensions of these ideas, but this short summary captures the common core of the teleosemantic proposals of such philosophers as Millikan (1984), Dretske (1995), Papineau (1987), and Neander (2017).

Selectionist teleosemantics is attractive for a number of reasons, one of which is that there is impressive support for the etiological theory of functions: it really does seem to capture what biologists have in mind when they engage in function talk (Neander 2017). Another reason is that, as we observed in Chapter 1, selectionist teleosemantics promises to explain the normativity of representation. Representations can *mis*represent the items they refer to, and they can also *mis*represent the world as containing items that it lacks, such as unicorns and phlogiston. That is, there is such a thing as representational error. But to describe a representation as erroneous is to evaluate it normatively, classifying it as a kind of failure. Selectionist teleosemantics maintains that it is possible to reduce this sort of normativity to malfunctioning—the failure of states or mechanisms to perform the functions that natural selection has conferred on them.

This brings us to the Swampman objection. Swampman is a human-like creature that comes into existence as a result of quantum fluctuations in swamp gas (Davidson 1987). He is endowed with a brain, internal organs, and sensory systems just like ours, and, as a result, he engages in the same sort of perceptual processing as we do, and he also engages in the same perceptually directed behaviors. As a result of these similarities, we are strongly disposed to think of Swampman as perceiving the external environment. But Swampman's perceptual states cannot be explained in selectionist terms. After all, Swampman is not a member of a species with an evolutionary history. Hence, it seems that he is a counterexample to selectionist accounts of perception. Further, since Swampman's brain is just like ours, there is a very strong inclination to think that Swampman must enjoy perceptual phenomenology that is just like ours, and must also be aware of that phenomenology. But this would be impossible if awareness of phenomenology was an essentially biological affair.

Although this objection seems compelling at first sight, reflection shows that it is actually quite misguided. The main point to notice is that it should be science, and not the philosophical imagination, that determines our beliefs about the essential natures of natural phenomena, including perceptual experience and perceptual phenomenology. Our ordinary concepts of natural phenomena often track what science reveals to be natural kinds, but these tracking relationships tend to be quite loose. Accordingly, while the philosophical imagination is a useful tool when one is concerned to chart the boundaries of the sets of actual and possible objects to which our ordinary concepts apply, it can be a poor guide to the boundaries of a natural kind that is imperfectly tracked by the concept. It is

science, and not a priori reflection, that can chart the latter boundaries. Moreover, it is the latter boundaries that should matter when we are concerned to map the essential structure of reality. Metaphysics should be the story of what there is, not the story of how our everyday concepts represent what there is.

In a lovely paper, Karen Neander illustrates these points vividly by asking us to imagine Swampcow, a creature that is exactly like a cow except that it comes into existence as a result of quantum fluctuations in swamp gas. Now contemporary evolutionary biology individuates species in terms of shared phylogenetic ancestry; that is, the boundaries of a species are determined by common descent. Because of this, Swampcow definitely does not count as a cow according to contemporary scientific standards. Could a philosopher sensibly maintain that these standards have been overturned because she has been able to *imagine* Swampcow? Neander playfully imagines a philosopher rushing into a meeting of evolutionary biologists to announce that she had imagined a counterexample to their evolutionary theory of what makes a cow a cow. What sort of reception would the philosopher receive? (Neander 1996)

Empirical theories cannot be challenged by imagined counterexamples. A related but different objection is that a counterfactual scenario involving Swampman would not necessarily play out in the way that critics of teleosemantics suppose. To see this, assume with me that science gives us excellent reason to believe that all awareness, including awareness of phenomenology, constitutively involves representations. Assume also that the best explanation of perceptual representations is selectionist, where by "the best explanation" I mean the explanation that is most in keeping with scientific talk of biological functions, including talk of informational functions. Suppose now, counterfactually, that scientists are confronted with Swampman. Would they credit him with perceptual experience? They shouldn't! It follows from the assumptions that Swampman would not be aware of visual phenomenology. To be sure, he would be in states that would covary with Thouless properties, but this sort of covariation, taken by itself, would not amount to representation and would not support awareness.

On the other side of this issue are intuitions to the effect that Swampman would be credited with phenomenal awareness because he is physically similar to ourselves—that is, similar to ourselves in terms of intrinsic physical properties and the causal powers that supervene on such properties. Many philosophers have been impressed by these intuitions, but I see them as raw and untutored, and therefore untrustworthy, because they are not grounded in an account of what awareness of phenomenology consists in. Epistemologically, they are on a par with the intuition that Swampcow "must" be a true cow because it is intrinsically similar to true cows.

To be sure, scientists might eventually be moved by the similarities between Swampman and ourselves to extend talk of awareness, perception, experience, and so on to him. It is assumed, after all, that Swampman is built out of the same material as ourselves, that he talks and behaves in the same way as we do, and that his internal states and processes are isomorphic to ours. It is also assumed that his

higher-level states encode the same information as our corresponding states. In general, Swampman shares our information processing capacities. But it does not follow that if scientists were moved by such considerations to expand the scope of talk of awareness, perception, experience, and so on, they would still be using concepts that can be appropriately applied to us. On the contrary, as the story about Swampcow indicates, the scientists' concepts would be new, even if the words used to express them were the same. The scientists would in effect be recognizing a new species of creature, defined in terms of material composition, causal powers, and encoded information, not in terms of biological functions and genuine representations.

A final word. Some philosophers would have concerns about this treatment of Swampman on the grounds that it does not do justice to his moral standing. (According to Papineau 2001, this concern originated with Eilert Sundt-Olhson.) Once we came to appreciate all of the similarities between Swampman and ourselves, it is maintained, we would surely treat him as having basic human rights. We would be quite confident, for example, that it would be morally wrong to eat him (Williams 2019). This seems right, but it can be fully explained, I suggest, by a tendency toward inclusiveness that is built into our moral principles. This tendency is not strong. Indeed, it is easily overridden by the exigencies of life, and even by self-serving biases of various kinds. But it does exist, and when we are presented with powerful evidence of similarity, and have no reason to ignore that evidence, it can take center stage. In the case of Swampman, the evidence would be overwhelming, in part because he was built of the same material as ourselves, in part because he was causally isomorphic to us, and in part because his perceptual and mental states encoded the same information as ours. But it does not follow that Swampman would be a human being, and, by the same token, it does not follow that biologists and psychologists should attribute the same biological functions to his organs, states, and processes as they attribute to the corresponding entities in us. The tendency toward inclusiveness that is built into moral categories is not matched by any tendencies in scientific categories. The purposes that classification serves are quite different in the two cases.

VI. Conclusion

So far, this chapter has been principally concerned to make a case for four theses. One is that representationalism provides a strong reason for recognizing that awareness of perceptual qualia is governed by an appearance/reality distinction. Another is that this distinction undercuts the motivation for property dualism. A third thesis is that representational accounts of phenomenal awareness are well motivated. And the fourth is that mere thought experiments cannot call well confirmed scientific hypotheses about the fundamental natures of things into question.

APPENDIX I

The Phenomenology of Conscious Thought

The view I wish to recommend is that this type of phenomenology (hereafter *cognitive phenomenology*) is exhaustively constituted by articulatory sensations and two forms of imagery. More specifically, the view is that the phenomenology of some thoughts consists largely of articulatory sensations in the mouth and neighboring regions, the phenomenology of other thoughts is constituted by visual imagery, and the remainder of cognitive phenomenology is constituted by auditory imagery. I believe that you will very likely agree that articulatory sensations and visual imagery constitute a significant portion of the phenomenology of thought, and I hope you will be convinced by arguments elsewhere in the book that the phenomenology of bodily sensations and the phenomenology of visual imagery can be explained in representationalist terms. Hence, the appendix will be concerned exclusively with the portion of cognitive phenomenology that can't be reduced to either of these other two phenomenologies—that is, cognitive phenomenology that is non-articulatory and non-visual. As noted, I think it plausible that this third form of phenomenology can be explained as auditory imagery. My goal will be to make a case for this belief. More specifically, my goal is to make a case for the view that the phenomenology of thought consists in auditory *linguistic* imagery. Call this the *auditory-linguistic* (*A-L*) *claim*. I will cite five kinds of evidence for it.

First, the A-L claim is supported by a range of dual-task experiments in which a linguistic activity like verbal shadowing (immediately echoing verbal material that is presented to the ear) disrupts high-level cognitive tasks that clearly require thought. In a typical study, Newton and de Villers (2007) found that shadowing English sentences interfered with reasoning about false beliefs, but that there was no such interference when the verbal shadowing task was replaced by an equally demanding one requiring subjects to shadow a rhythm by tapping. (See Hermer-Vazquez *et al.* 1999 (but cf. Ratliff and Newcombe 2008 for important qualifications); Winawer *et al.* 2007; Lupyan 2009; Frank *et al.* 2012.) Note that shadowing requires attentive tracking of auditory signals, and therefore likely interferes significantly with auditory imagery.

Second, many bilinguals report palpable differences between, say, thinking in Russian and thinking in English, and they are often able to identify types of thinking (for example, mathematical calculations) that are more likely to be conducted in one of their languages than the other (Guerrrero 2005, Dewaele, 2011). This suggests that thinking in Russian is associated with a different phenomenology than thinking in English. But the two phenomenologies must have something to do with language. Moreover, it is natural to suppose that the phenomenologies involve auditory imagery. The data are also suggestive of articulatory imagery, but it is very implausible that articulatory imagery is always involved in the thinking of bilinguals.

Third, the A-L claim is supported by the testimony of subjects in experiments conducted by the innovative psychologist Russell Hurlburt. Hurlburt has assembled a large body of relevant data by using a technique that he calls *Descriptive Experience Sampling*. (In a DES experiment, the subject carries a pager that produces beeps at random intervals, and takes notes on what is before their mind at the moments when the pager sounds. The experimenter discusses the notes with subjects at a later time, with a view to eliciting clarifications. (See, e.g., Hurlburt and Schwitzgebel (2007).) Hurlburt's subjects report that their thoughts are often associated with auditory imagery. They describe them as instances of "inner hearing."

But this isn't the whole story. Hurlburt's subjects also say that their thoughts sometimes take the form of inner speech, and they contrast these thoughts with the ones that they classify as examples of inner hearing. Does this pose a problem for the A-L claim? I don't think so. When Hurlburt's subjects report inner speech, they are probably often alluding to the kind of thought that involves articulatory sensations. There is no doubt that thought can take the form of subvocal articulations, but the phenomenology of that sort of thought is a special form of sensational phenomenology generally, and it is therefore treated in Chapter 6, albeit implicitly. (Much of what I say there about the phenomenology of pain can be generalized to apply to all sensational phenomenologies.) Reports of inner speech can perhaps also be explained in terms of the efference trace model of inner hearing, which I discuss below.

In addition to episodes of inner hearing and inner speech, Hurlburt's subjects claim to have episodes of "unsymbolized thinking"—that is, "thoughts that have clear, differentiated content but no discernible features that 'carry' that content: no images, no words, no other kinds of symbols" (Hurlburt and Schwitzgebel 2007: 266). Again, I see no challenge to the A-L claim. Reports of unsymbolized thinking certainly suggest that there are thoughts that are not accompanied by auditory imagery, but they also suggest that the thoughts in question have no associated phenomenology of *any* kind. The subjects indicate that they are aware of contents, period. Testimony of this sort is perfectly compatible with the A-L claim. The A-L claim maintains that the phenomenology of thought is constituted by auditory imagery, not that all thoughts are accompanied by phenomenology.

Fourth, there is beginning to be fMRI evidence for the linguistic basis of the phenomenology of thought. The brains of subjects with DES training have been scanned with a view to determining whether linguistic centers such as Broca's area and Wernicke's area are active when subjects report episodes of inner speech and inner hearing. (See Shergill, et al. 2001, Kühn et al. 2014, Hurlburt et al. 2016, Loevenbruck et al. 2018, and Langland-Hassan 2020. Broca's and Wernicke's areas are concerned, respectively, with the production of speech and the auditory perception of speech.) The evidence is so far quite limited, and has more bearing on inner speech than inner hearing, but it does point toward the engagement of linguistic centers when conscious thought is occurring. I should add that the issues here are delicate. Suppose it is true that cognitive phenomenology consists of auditory linguistic phenomenology—that is, of auditory images of words, phrases, and sentences. This imagery must have causes, and it is plausible that the causes are generative processes of sentence construction that are more likely to occur in Broca's area than in Wernicke's area. Accordingly, it is to be expected that scanner evidence concerning inner hearing will be deeply entangled with evidence that might be thought to point to inner speech.

Fifth, there is a plausible hypothesis about the speech production process that predicts the occurrence of auditory linguistic imagery. This hypothesis is part of the more general theory that the processes responsible for producing action plans broadcast those plans to perceptual areas. These broadcasts are known as *efference traces*. It is believed that, among other things, they help agents to predict the sensory consequences of their own actions. In the following passage, Peter Carruthers argues that auditory linguistic imagery can be seen as composed of efference traces of speech production. (By "inner speech" Carruthers pretty clearly means what I have been calling "inner hearing.")

> There is also broad agreement about *how* motor planning gives rise to inner speech. Motor schemata are selected and activated . . . , issuing in an "efference copy" or "corollary discharge" of those instructions while downstream activation of the facial and throat musculature are suppressed. The efference copy is used to create a sensory representation of what the speech act in question would have sounded like had it been

> carried through to completion (Tian & Poeppel, 2010; Corley et al., 2011; Scott, 2013; Scott et al., 2013). Indeed, we know from recent work using brain-scanning of various kinds that the process proceeds in two stages, with motor activations first creating a somatosensory representation of what the movements in question would feel like, which then creates an auditory representation of what the resulting speech would sound like (Tian & Poeppel, 2013). The process begins with activity in Broca's area (the classic speech production/speech-planning area), which then activates the relevant regions of primary motor cortex. This then causes activity in somatosensory regions of parietal cortex, before activity is produced in abstract auditory regions of the superior temporal gyrus and superior temporal sulcus. (Carruthers 2018, 33)

The details of this proposal are subject to revision, but the underlying idea that there are auditory efference traces of plans for speaking is part of a more general idea that has proved useful elsewhere in cognitive neuroscience. (See, e.g., Fleming 2021.)

Arguably, the fact that auditory efference traces are produced by processes like the ones that generate overt speech helps to explain the testimony of Hurlburt's subjects concerning inner speech.

The A-L claim has recently been endorsed and convincingly defended by Jesse Prinz (Prinz 2012, chapter 5). Here is his version of the idea. (The context strongly suggests that Prinz is using "inner speech" to mean what I mean by "inner hearing.")

> I suspect that verbal imagery can explain every instance of conscious thought that cannot be accounted for by appeal to [perceptual] images of the contents of our thoughts (what I will sometimes call simulations). Inner speech is incessant and implicated in many aspects of cognition . . . It is also known to be underwritten by the brain mechanisms involved in speech production and, importantly, perception. . . . Given the conspicuous presence of silent speech in the stream of consciousness, it seems likely that we often come to know what we are thinking by hearing inner statements of the sentences that we would use to express our thoughts. (Prinz 2012, 159)

The arguments I have given for the A-L claim are different from, and are meant to complement, the ones offered by Prinz. Alex Byrne and Peter Langland-Hassan are also advocates of the A-L claim, and they too give substantial independent arguments (Byrne 2018, Langland-Hassan 2018). The reader is referred to the cited works by these authors for further discussion.

This brings me to a concern about the A-L hypothesis that I have not yet fully addressed. As we observed in considering Hurlburt, it seems that thinking is fundamentally different than such paradigmatic cases of inner hearing as songs running through one's head. (Speeches can also run through one's head—e.g., the Gettysburg Address, delivered in the intonations of Daniel Day-Lewis. (Day-Lewis played Lincoln in a memorable movie.)) In paradigm cases, one is, as it were, conscious of tokens of particular words. It is very much as if one was actually hearing the words. But informal surveys have convinced me that this is not how most of us experience thought. In thought, there is no imagery that could be described as imagery of a sequence of phones. (A phone is a concrete token of a phoneme, much as a particular mound of graphite on a piece of paper might be a token of an Arabic numeral.) If it is auditory imagery that we experience in thinking, the imagery is more abstract than that. Is this compatible with the A-L hypothesis? One might be inclined to give a negative answer.

There is, however, a response to this concern. When we are perceiving overt speech, there are several levels of representation, connected by inference. At a low level, what we perceive is sequences of phones. But at a slightly higher and more abstract level, the

sequences of phones are represented as sequences of phonemes; and at a still higher level, the sequences of phonemes are represented as divided into groups that correspond to words. There is also a level at which sequences of words are represented as having syntactic structures, and eventually, a level occupied by representations of the meanings of whole sentences. Now it seems that each of these levels of auditory representation could be activated without activating lower levels. For example, it seems possible for a representation at the phonemic level to occur without its being triggered by a lower-level representation of a sequence of phones. The representation would be caused by corollary discharge, rather than an external stream of speech, and it would have a (false) content roughly of the form, *there is currently a sequence of phones F1, F2, . . . , Fn such that F1, 2, . . . , and Fn are tokens, respectively, of the phonemes P1, P2, . . . , Pn*. Now if all of this was true, then it might also be true that efference traces of plans to speak take the form of representations at the phonemic level. Or they might take the form of representations at a still higher level. In either case, the auditory representations involved in thought would be more abstract than the representations that occur when a song runs through one's head, and, a fortiori, more abstract than the ones that occur when one hears a stream of overt speech.

To summarize, there are a number of reasons for accepting the A-L claim. Moreover, a pressing objection to the thesis may admit of an answer. In view of all this, there are grounds for optimism about the prospects of representationalist accounts of cognitive phenomenology. (Not that this should be seen as a settled question: exploration of the empirical issues is still at a very early stage. The interested reader should consult Langland-Hassan 2020 for an extensive and balanced review of the recent literature.)

APPENDIX II

The Phenomenology of Moods and Emotions

As is the case with cognitive phenomenology, work on the phenomenology of moods and emotions is still at a primitive stage. Even though affective neuroscience and affective psychology have made enormous strides in the last twenty-five years, they are not nearly as advanced as perception science. Moreover, the phenomenology of affective states seems more elusive than perceptual phenomenology. Introspection has less to tell us, and what it does tell us seems less trustworthy (Schwitzgebel 2008, Section IV, Schooler and Maus 2010). The proposals I will make in this appendix will, I hope, make it plausible that there can be a representationalist account of affective phenomenology, but they will for the most part be vague and leave much room for future work.

Before we consider affective phenomenology, we should recall that affective states have large and complicated causal or functional roles. Consider, for example, the following functional profile of deep depression:

> Physiological symptoms include sleep disturbance, most often insomnia with early morning awakening but occasionally excessive sleeping; loss of appetite and weight loss, but occasionally excessive eating; and decreased energy. Behaviorally, some depressed patients exhibit slowed motor movements, described as psychomotor retardation, whereas others can be extremely agitated. Cognitive symptoms are evident in both the content of thoughts (hopelessness, thoughts of worthlessness and of guilt, suicidal thoughts and urges) and in cognitive processes (difficulty concentrating, slow thinking, and poor memory). (Hyman and Cohen 2013, 1403–4)

When we reflect on other affective states in relation to this portrayal of deep depression, we appreciate that most of them resemble depression in that they too have physiological, behavioral, and cognitive dimensions, though of course these dimensions do not usually involve such major departures from normality.

Turning now to the phenomenological aspects of affective states, it appears to be widely held that it has three different dimensions—bodily sensations such as the ones arising from an accelerated heartbeat, higher-level qualia such as the ones associated with experiencing a bear as threatening or rotten food as disgusting, and valenced aspects of rewarding and aversive stimuli such as the pleasantness of the taste of ice cream. I will discuss each of these types of affective phenomenology in turn, with a view to urging that they can all be understood in representationalist terms.

William James famously believed that emotions can be analyzed without remainder into impressions of bodily occurrences of various kinds, including especially muscle tensions, heart rate, depth and frequency of breathing, and visceral processes (James 1884). This view cannot be correct because it ignores the cognitive and behavioral dimensions of emotions that we noted a moment ago. That is, it cannot be correct as a theory of emotional *states*. But it could still be true as a theory of emotional *phenomenology*. More specifically, it could still be true in the version which claims that all emotional phenomenology is explainable in terms of mental representations of bodily states and processes. James's arguments lend some plausibility to this second view, the *somatic theory* of affective phenomenology, and his arguments have been broadened and strengthened considerably in recent years by the work of Antonio Damasio (Damasio 1999, Damasio 2003, Damasio 2021). Here is an illustrative set of arguments:

> There's support for the idea—three of my favorites are that (a) forcing depressed people to smile makes them feel better; (b) instructing people to take on a more "dominant" posture makes them more so (lowers stress hormone levels); and (c) muscle relaxants decrease anxiety. . . . (Sapolsky 2017, 91)

Reasons of this kind can be quite persuasive. And there are many of them. As James noted, "If the reader has never paid attention to this matter, he will be both interested and astonished to learn how many different local bodily feelings he can detect in himself as characteristic of his various emotional moods" (James 1884, 193). No wonder that some people have devoted considerable energy to defending the somatic theory. (Chapter 7 of Hill 2009 is an extended defense. The present paragraph abridges that earlier discussion but the appendix hereafter proceeds in a different direction.)

I will argue in the next chapter that somatic qualia can be accommodated by representationalism.

There is no doubt that the somatic theory can account for an enormous amount of affective phenomenology, but can it be right in claiming to explain all of it? It appears from the literature that many people believe that there are other, non-bodily aspects of phenomenology that it fails to acknowledge. These writers do not often give reasons for rejecting the view, but some can be constructed.

The Stoicism argument: Consider an "external Stoic" who has learned to keep all but her most extreme emotions hidden from external observers. She never adopts postures or facial expressions that express her emotions, and she has trained herself not to engage in behaviors that would reveal them. She has even managed to control her breathing to a large degree, so that someone watching her chest could not easily detect emotional states. Couldn't such a person still have feelings of fear or sorrow? It seems that the natural, intuitive answer is "yes." No doubt some Jamesian somatic qualia would remain. For example,

her heartbeat would probably be at least mildly accelerated in moments of fear. It seems unlikely, however, that the remaining somatic qualia would be able to uniquely specify the way she felt. A mildly accelerated heartbeat is common to a number of different emotions, as is a slightly elevated breathing rate.

The relaxation argument: Consider an episode of deep meditation, or of lovemaking, or of swimming in the surf, and the profound sense of relaxation that comes afterwards. Your heart isn't active in a way that is indicative of emotion; your breathing is neutral; and your facial muscles are not tightened to form a smile or a frown. Nor are your thoughts racing or sluggish. You are simply lying somewhere—perhaps on a bed or on the sand. But still, it seems, even though you are in deep repose, your consciousness might be tinged by joy or sorrow. Perhaps you experience a soft glow of joy because you feel that your recent lovemaking has added a new depth to an old relationship, or you experience a trace of melancholy because you know it will be months till you see your partner again. It is hard to see how these feelings could consist of bodily sensations.

The motivation argument: We believe that feelings of joy, sadness, fear, anger, surprise, and disgust are motivating: they support dispositions to engage in various types of cognitive and overt activity. Moreover, it is known that these states are normally accompanied by facial expressions that serve as signatures: the expressions uniquely specify the underlying affective states. Now suppose that the somatic theory of affective phenomenology is true. It follows that representations of the muscle tensions involved in these facial expressions must motivate, or at least partially motivate, the cognitive and behavioral dispositions that normally accompany the foregoing six emotions. But is it credible that a mere representation of a facial expression can be motivating, sometimes driving agents to extraordinary lengths?

These arguments are less than decisive, but they seem to provide reasonably strong motivation for thinking that affective phenomenology has non-somatic components.

A moment ago I cited the idea that there are three dimensions of affective phenomenology. We have been considering the first of these three dimensions, noticing, among other things, that it cannot stand alone as an explanation of affective phenomenology. I turn now to the second type of affective phenomenology.

If Jamesian somatic qualia do not exhaust the phenomenologies that are characteristic of emotions, the remaining qualia must be either internal to the subject and mental or external to the subject.

The idea of internal emotional phenomenology is suggested by introspection and folk psychology. Thus, introspection suggests that there is something profoundly internal about feeling afraid or angry or cheerful, and folk psychology confirms this introspective impression by classifying emotional phenomenologies as feelings and feelings as mental states. Moreover, as I will soon explain at some length, it is often difficult to see what the objective correlate of an emotion might be. Moods, for example, can seem to be independent of external conditions. On the other hand, it is a salient fact that many emotions do have external correlates. It is natural to think that my fear of the bear ahead of me on the path involves a representation to the effect that the bear has the teleological property *posing a threat*, and/or the response-dependent property *frightening*. By the same token, it is tempting to suppose that the phenomenology associated with my fear of the bear can be explained in terms of a representation with one or both of these contents.

I will eventually propose an externalist account of emotional phenomenology, but before getting to that, I want to expand on the point that many emotions seem to lack external correlates. It threatens to sink externalist theories of emotional phenomenology before they are unfurled. We need a deeper understanding of this issue before we can proceed.

As noted, there are no obvious external correlates for the emotions that count as moods, such as generalized anxiety, generalized optimism, despair, boredom, and a sense of calm. It is sometimes suggested, however, that emotions of this sort actually have external correlates of a special and unusual kind. It has been proposed, for example, that despair has a content involving the whole World—perhaps the content *the World contains nothing of value (that is, nothing worth pursuing or even attending to)*. But despair is an emotion that seems cognitively impenetrable, and therefore one that is unlikely to represent the whole World, given that our representational grasp on the World and its properties is conceptual and cognitive. To be sure, there are often judgments censuring the World as a whole that accompany basic despair ("How weary, stale, flat, and unprofitable seem to me all the uses of this world!"), but such judgments are best seen as effects of the emotion, not as causes or constituents of it.

Could it be that the content of despair is the more restricted state of affairs *my immediate environment contains nothing of value*? No, if that were true, travel would be the obvious and standard cure for despair. But, as the poet C. P. Cavafy observed in "The City," it would be foolish to hope for more than temporary relief from despair by sailing for foreign shores. What about the idea that the content of despair is the state of affairs *there is nothing of any value anywhere*? Like the first suggestion, this one makes the mistake of attributing high-level cognitive content to a lower-level emotional state. Concepts are required to express negative existential content. Similar reflections apply to generalized anxiety, boredom, and other paradigm cases of moods.

Joy differs from these states in that it typically has a manifest external cause (the death of a tyrant, an apt birthday present, an audience with the Pope), but it is similar in that it often lacks a *continuing* external correlate. You rejoice when you first see a new baby, but your joy may persist for hours or days after that encounter, even though the birth and the baby have long since ceased to dominate your thoughts. Or it might be that you awake in a joyous frame of mind. No doubt there is a cause of your emotional state—perhaps you received good news in the middle of the night. But a cause of joy isn't the same as an object of joy, an item on which joy is directed. There need not be anything that is a continuing object of your joy, though of course your joy may bestow a radiance on things that you encounter or think about as the day proceeds. ("A mouse is eating the cheese set out for lunch—how adorable it is!" "I'm thinking of the Fibonacci sequence—a gorgeous object!") Similarly, an external item (a quarrel, a missed train) may precipitate an angry mood that persists throughout the day, casting a pall over all that happens, even though one would be hard put to identify a continuing object of one's anger. In general, precipitating causes of emotions need not continue as their targets.

In view of these considerations, it is initially hard to see how the phenomenology of moods could consist in representations of external states of affairs. It is tempting to suppose that it is irreducibly mental in character. Another reason for thinking that emotional phenomenology is internal is that even emotions like fear, which are plausibly regarded as having external correlates in most cases, can occur in diffuse and unfocused forms. Thus, you might experience fear when a fortune teller predicts that you will soon meet with an unspecified disaster, or when a traffic accident makes you aware of your general vulnerability to untoward forces, or when you are lying in bed late at night, alone and sleepless, and the contingency of your existence and all that you care about overwhelms you. In all these cases, certain conditions precipitate a state of hypervigilance and generalized anxiety that is experienced as fear. But the fear isn't directed on a specific entity.

It might be suggested that the phenomenon I'm describing isn't a continuing state of unfocused fear, but rather a disposition to become newly afraid on encountering certain

specific stimuli. Reflection, shows, however, that this suggestion does not do justice to the phenomenon. Imagine a recently injured human being skulking through a forest, constantly glancing warily from side to side. We can suppose that the person isn't even sure of what *kinds* of threat the forest might present, and also that the person is not afraid of the forest itself. (It hasn't occurred to the person that, for example, the whole forest might spontaneously catch fire, spouting flames everywhere.)

This sort of generalized anxiety or wariness seems to have its roots in the dispositions of members of more primitive species. Consider the following passage from Peter Godfrey-Smith's captivating book, *Metazoa* (p. 212):

> Terry Walters of the University of Texas has worked for many years on a fear-like state called *nociceptive sensitization*. This is a heightened sensitivity after damage that alters the animal's response to various *other* choices and stimuli. The animal seems to have an overall wariness, a caution that can last for hours, days, or weeks, depending on what is the case.

Godfrey-Smith goes on to note that more primitive animals also display generalized, unfocused attitudes of other kinds (pp. 212–13):

> A similar emotion-like state was seen by Melissa Bateson and her colleagues in honeybees. They found that if a bee is shaken, it induced a kind of pessimism: the bee became inclined to treat ambiguous stimuli (midway between what had been good in earlier experience and what had been bad) in a pessimistic manner, assuming the worst. On the positive side, bees can also be induced to be upbeat or optimistic: Cwyn Solvi and her collaborators in Lars Chittka's laboratory showed that an unexpected reward has an effect on bumblebees that is the flip side of the pessimistic mood that Bateson's experiment produced. Good moods are also seen in fish.

When I think about these animal studies, I worry that they might cause a utilitarian reader to feel guilty, since they point to many heretofore neglected ways to promote the general welfare.

To summarize: it initially appears that there are strong reasons, having to do with a lack of external correlates, for regarding some aspects of emotional phenomenology as internal and mental.

When we reflect, however, we see another side to this issue. Moods and unfocused emotions may not attribute properties like *posing a threat* to specific external *objects*, but they could still represent external *conditions*, in the way that sentences like "It's raining," "It's foggy," "It's cold," and "It's night" represent environmental conditions without attributing them to specific objects or assigning them to specific regions. To illustrate, despair might have a teleological content like *it's completely lacking in rewards!*, joy might have a content like *it's beneficent!*, and unfocused fear might have a content like *it threatens!*, where the "it"s are to be understood as non-referential. Alternatively, we can suppose that these emotions have the contents *unrewarding!*, *beneficence is manifested!*, and *threatening!* In other words, the idea is that the emotions that lack specific external correlates are constituted by representations of external *circumstances* that are vaguely located. The difference between these states and emotions that are focused on specific external objects is that in the latter case, representations that express properties of objects are employed, and are bound to perceptual representations of specific external objects. Thus, in the case in which I see a bear on the path ahead of me, I form a complex representation to the effect that the bear poses a threat, where *poses a threat* is a property attributed to the bear. The contents of representations of conditions and representations of properties of objects are closely

related, and it may even be that there is only one class of representations that function in different ways in different contexts. (Cf. Quine's discussion in *Word and Object* of how the word "gavagai" might function as a stand-alone "occasion sentence," and also as a term that serves as a proper part of sentences (Quine 1960).) The bottom line, though, is that there are two kinds of emotion that possess different contents, albeit contents that can be very closely related.

As we noticed in passing earlier, there are two views about the external contents of states of fear. One is that they represent external entities as having the teleological property *posing a threat*, and the other is that they represent external entities as having the response-dependent property *frightening*—that is, the property *causing (my) fear*. I prefer the former view because it is more in keeping with a plausible theory about the nature of fear. According to the theory I have in mind, fear that is focused on an external object O has four principal components: a range of Jamesian somatic qualia, a perception of O, a representation of O as having an additional property P, where P explains the non-Jamesian qualia of fear, and an attitude toward O, where the attitude is responsible for motivating mental and physical activities that are ways of responding to O. The question before us now is, "What is P?" Is it the property *posing a threat* or the property *causing my fear*? It must be the former property. First, if perceptual representation involves causation, then we would have to say that the representation of P is caused by the property *causing fear* (where fear includes the representation of P). But nothing could cause a state X in virtue of having the property *causing X*. Causal properties aren't causes. Second, why would an agent need to represent all of the components of a state of fear, including Jamesian qualia and perceptual awareness of O? It suffices that the agent is somatically representing Jamesian qualia and perceptually representing O. Representing the representations of the qualia and O is *de trop*, as is representing a representation of P, whatever P may be. What the agent needs to know about is the threatening character of the perceived object, not these features of its own subjective state. (Of course, the agent will represent the *behaviors* that are caused by fear at a later stage, in order to estimate their margin of success or failure, but this sort of self-representation occurs after the onset of fear and is not part of fear itself.)

Similar considerations apply to the representational contents of other emotions. Human beings need to represent the property of being putrid or foul. We should suppose that it is that property, rather than the property *disgusting* (*causing disgust*) that is represented by disgust. Equally, humans need to represent the property of being obstructive or encroaching, so we should suppose that it is that complex property that is represented by anger. There is no comparable need to represent a property like *vexing* (*causing vexation*) or *infuriating* (*causing anger*).

To develop the present proposal a bit further, it is meant to imply that certain states of the limbic system are representations on par with the more widely recognized representations in the various perceptual systems. These limbic states stand for external conditions and/or properties of specific external entities, and it is this representational role that explains the non-Jamesian aspects of the phenomenology of emotions. More specifically, emotional qualia are the conditions and the properties that these limbic representations signify. To give an example, on the present proposal, the phenomenology of fear can be explained in terms of a certain pattern of activity in the limbic system—call it *activity of type F*—that has the status of a representation. The phenomenology is constituted by the teleological contents *threatening!* and *poses a threat*, the former being an environmental condition and the latter being a property of external entities. Activity of type F can represent these external contents because it has covaried with them historically, and because this covariation has conferred a selective advantage on the creatures it involves (Millikan

1989a, Dretske 1995, Neander 2017, Shea 2018). The limbic system is known to be the primary cause of the physiological, behavioral, and cognitive signatures of fear (LeDoux and Damasio 2013), so it should not surprise us if it plays a causal role in generating in fearful phenomenology.

What about the fact that fearful phenomenology is conscious? Can the present proposal honor this fact, given that the limbic system is largely subcortical? Yes, provided that we make an adjustment. It is known that multiple parts of the limbic system are connected with cortical areas. For example, the amygdala is richly and reciprocally connected with the orbitofrontal cortex. (See, e. g., Rolls 2019.) Hence, there is a basis for saying that the representations that are responsible for fearful phenomenology span limbic and cortical sites. And we *should* say that, since there are good reasons for thinking the consciousness requires cortical activity.

If the present account of the phenomenology of emotions seems to leave something out, I suggest it is because it is hard to take on board the moral of the earlier discussion of the unbridgeable gulf argument. I urged there that intuitions about the irreducibility of qualia arise because folk psychology recognizes no distinction between the appearance of qualia and the corresponding reality. If, for example, the quale *blue* seems metaphysically sui generis, it is because we can see no way in which the quale might turn out to be an appearance of something else—something that we grasp in different ways in other contexts. The same point applies to qualia like *posing a threat* and *actually or potentially beneficial*. Like all other qualia, they can seem to be metaphysically sui generis. But that is because folk psychology provides no way for thinking of them as appearances. Folk psychology gets many things right, but parts of it are dark prisons that block out all light and restrict thinking to circular movement.

A number of emotions, such as fear and anger, have an "electricity" or "dynamism" that accounts must explain, and this characteristic seems to be bound up with, though not necessarily part of, their phenomenology. When we experience their phenomenology, we are simultaneously motivated to act in certain ways, or at least motivated to consider various courses of action. The present proposal can accommodate this aspect of phenomenology. Consider fear again. As is well known, there is excellent reason to think that activity in the amygdala plays a key role in producing the psychological and behavioral signatures of fear. More specifically, when it occurs at a certain level, such activity automatically and inflexibly induces the hypothalamus to trigger the release of hormones needed for a fight or flight response. This inflexible, sub-cognitive causal power of the amygdala is a principal cause of the "electricity" of fear. It is bound up with the phenomenology of fear because, according to the present proposal, activity in the amygdala also plays a role in producing representations of properties like *posing a threat*. Indeed, it may be that both the motivational role and the phenomenology of fear are due to the very same type of amygdaloid activity—the type I earlier called activity of type F. In general, the phenomenology of fear appears to be inextricably bound up with behavioral and cognitive reactions because they have a common source and therefore have a strong tendency to co-occur.

This story is of course largely conjectural, but it is probably the account that makes best sense of the data that are currently available.

At the outset of this appendix, we took note of the idea that there are three types of affective phenomenology—somatic qualia, high-level, non-somatic qualia, and the valenced aspects of rewarding and aversive stimuli. We now have accounts of the first two types of phenomenology. It is time to turn our attention to the idea that positive and negative valence have phenomenologies. In considering this topic, I will focus on the nature of reward. I will consider the corresponding idea concerning aversive stimuli in the next chapter.

The view I wish to defend is that there is no phenomenology of reward *per se*. Consider what happens when you find ice cream delicious. On the view I wish to defend, your experience of the ice cream consists in attentively representing certain "factual" or "natural" properties of the ice cream, and, in particular, its taste and olfactory bouquet. There is no separate factual property of the ice cream in which its deliciousness might consist. To put it another way, deliciousness is not one of the properties of the ice cream that we are aware of when we experience it as delicious. Deliciousness is not itself a property of the ice cream nor an object of awareness. Instead, it is an attitude we take toward the ice cream, an attitude that moves us to continue eating it and perhaps order more.

The reason for this view is that finding ice cream delicious is *intrinsically motivating*. One cannot find ice cream delicious and not desire to savor it and continue to consume it. But this means that deliciousness is not a factual or natural property, because no factual property is intrinsically motivating. Only attitudes of various kinds—needs, desires, preferences, likings, etc.—are constitutively associated with motivation.

One might dispute this claim about factuality and the power to motivate, urging that normative properties like goodness (or, more pertinently, gustatory goodness) are both intrinsically motivating and factual. This is not the place to enter into a discussion of the objectivity of values. I will simply say that a naturalistic version of representationalism requires that represented properties be factual or natural in character, and it's hard to see how goodness could have the natural standing that a property must have in order to qualify as a representatum. At all events, it would be a significant cost of a theory of deliciousness if it was committed to the highly controversial claim that goodness is located within the natural order.

It seems we must say, then, that deliciousness is an attitude that agents take toward the factual properties of the objects that they find delicious. But is this compatible with the intuition that we are aware of objects as delicious? Isn't it true that when we find ice cream delicious we are aware of it as having *some* property over and above the factual properties that constitute its having the taste? It might seem that the present proposal has no way of answering that question.

My response is that we should not bow to this intuition. It seems no less plausible that when we are eating ice cream, we experience its taste and *like* that taste, where liking is an attitude toward the taste. Indeed, this is the way that we would normally describe such situations, suggesting that this way of conceiving of them is fundamental. Moreover, it is quite dubious that introspection can discern a property of ice cream over and above the factual properties that constitute its taste. Is it possible to focus one's attention on such a property? Is there anything left of our experience of ice cream after taste has been subtracted in the imagination? I suggest that the answers should be negative. But if that is right, we must conclude that careful, reflective introspection actually disconfirms the intuition. Perhaps it is less an intuition than an unsupported thought about what must be the case.

The idea that deliciousness of tastes consists in our liking them is confirmed by the experimental work of Kent Berridge and his associates (Kringelbach 2005, Smith et al. 2010, Kringelbach and Berridge 2015, Kringelbach and Berridge 2017, Nguyen et al. 2012), and also by related work by other investigators (e.g., Grabenhorst and Rolls 2011, Rolls 2019). In the interpretation of these studies that Berridge and his fellow scientists prefer, the positive valence of an experience is identified with liking the object of experience.

Of course it might be that states of liking, qua possessing various degrees of intensity, are themselves objects of representation. Indeed, this hypothesis is arguably required by

the theory of reward learning (Schultz et al. 1997, Schultz et al. 2000, Schultz 2006, Shea 2014). But I have been arguing that as far as introspection can determine, there is no conscious phenomenology associated with our response to valued stimuli other than the phenomenology involved in perceiving them. Moreover, Schultz and other contributors to the literature on reward learning seem to have little use for the idea that such learning involves a proprietary phenomenology.

6
A Quasi-Perceptualist Account of Pain Experience

I. Introduction

Like the previous chapter, this one is concerned with phenomenology. Taken out of context, that might seem to be a fool's errand, but, as we have seen, it is possible to make progress in this area by pursuing the hypothesis that phenomenology is determined by the contents of perceptual and interoceptive representations located at sites in the brain. I will be attempting to increase the plausibility of this hypothesis by supplying more details.

More specifically, I will argue that the foregoing account of perceptual qualia can be applied, with a few adjustments, to explain the phenomenology of pain and awareness of that phenomenology. I will also maintain that awareness of pain is sufficiently similar to paradigmatic forms of perceptual awareness that it is natural and appropriate to classify it as quasi-perceptual.

Discussions of pain are made difficult by the fact that the ordinary concept of pain is deeply embedded in a folk theory that is blind to key distinctions and verges on incoherence in places. Thus, the concept is used to keep track of importantly different phenomena that co-occur in normal cases, but can be separated by surgery and drugs, and are sometimes found to be dissociated in the clinic. Folk psychology does not provide us with the conceptual equipment needed to draw the relevant distinctions. Then, too, folk psychology suggests both that pain is mental, since it does not distinguish between the mental experience of pain and pain itself, and also that it is bodily, since it authorizes us to classify all pains with respect to bodily locations.

I will begin by saying a bit more about the flaws of the folk picture of pain and the attendant flaws of the associated commonsense concept. This will involve developing an alternative conceptual vocabulary that is free from these defects. After setting the stage in this way, I will go on to argue that when agents speak of awareness of pain in normal circumstances, what is actually occurring is awareness of activity in nociceptors that are spread across the periphery of the body, and are also buried in various internal components, including muscles, joints, and the viscera. This effort will provide the basis for an account of the phenomenology of pain that departs in numerous ways from the account that is embedded in folk psychology.

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0006

In Sections II–IV, I will assume that the phenomenology of pain is fairly simple, consisting only of determinate forms of the quale *sensory pain*. (The determinate forms are individuated by levels of intensity, locations, and such differentiae as *burning* and *pricking*.) In Section V I will consider the question of whether we should suppose that there is another pain quale, independent of all these, that might be called *hurting*. To anticipate, my answer to that question will be negative, but the relevant issues are complex and require extended examination. (The line of thought in Sections II-IV will combine and extend the ones in Hill 2014a, 2014b, and 2017.)

II. Fixing Ideas

All normal cases of pain involve the following five components:

1. A noxious stimulus, which may occur either on the surface of the body or in internal areas such as the muscles, joints, and viscera. The stimulus may involve outright damage or just a condition that portends damage or some other serious impairment of proper functioning. I will refer to instances of this component as *N-stimuli.* They include burns, punctures, tumors, and so on.

2. Activity in neurons that detect N-stimuli and are usually known as *nociceptors*. Nociceptors are principally of two types—A-δ fibers and C fibers. The former are comparatively large in diameter and correspondingly rapid in signal transmission. They are coated in a sheaf of white fatty tissue called *myelin*. The latter are small in diameter and correspondingly slow in transmission. They are unmyelinated. Both kinds of nociceptor are found at the surface of the body and also in the interior, and run from their initial points all the way to the spinal cord. I will refer to instances of activity in these nociceptors as *N-processes*. Please understand this to imply that the property *being an N-process* is the same as the property *being a case of activity in C-fibers or in A-δ fibers or both.* (Note that "nociceptor" picks out C-fibers and A-δ fibers in terms of their functional role, while "C-fiber" and "A-δ fiber" pick them out in terms of their intrinsic neurological characteristics. Although I will sometimes refer to C-fibers and A-δ fibers as nociceptors, in the interests of convenience, in speaking of N-processes, I will always mean to refer to neural phenomena.) There are also nociceptors of a third type—"silent" nociceptors that are found in the viscera (Basbaum 2021). Their contribution to pain seems to be comparatively small, so I will prescind from them here.

3. Events at the tips of A-δ fibers and C-fibers that initiate N-processes. I will call these phenomena *N-events*. They are closely related to N-processes: indeed, they are the initial parts of N-processes. But there is a significant difference here and it is highly desirable to keep it in mind. N-processes are extended in space, running, for example, from the biggest left toe all the way to the spinal cord. N-events are confined to sites of damage and potential damage.

4. A cortical response to N-events and N-processes. This response seems to be centered in the insula, though other cortical areas, such as the anterior cingulate cortex and somatosensory cortex, are generally involved. I will refer to instances of this response as *N-states* (Price 2000, Craig 2010, Craig 2015, Basbaum 2021).

5. An affective response to N-stimuli, N-events, and N-processes. This response involves activity in deeper, sub-cortical parts of the limbic system, such as the amygdala, but it also generally involves activity in cortical areas, even including the prefrontal cortex. I will refer to instances of the response as *N-responses.*

Although these five phenomena are intimately linked in normal cases of pain, they can come apart in various ways. Most notably, perhaps, N-responses and N-states can occur without being accompanied by N-stimuli or N-events or N-processes. This happens in cases of phantom limb pain. In this condition, subjects describe themselves as experiencing pain, and the rest of us acquiesce in their descriptions, recognizing first-person authority with respect to the use of the concept of pain; but since the relevant body part is missing, there is no peripheral damage or disorder corresponding to the experience, and no appropriately located activity in A-δ fibers and C fibers. Similarly, in many cases of chronic pain, there is no corresponding damage or disorder in the relevant part of the body, and the A-δ fibers and C fibers are silent. Further, amazingly, when N-stimuli, N-events, N-processes, and N-states occur in agents with a condition known as pain asymbolia, they are not accompanied by N-responses. Asymbolia patients describe themselves as experiencing pain, but they deny being bothered by their pains, and there are no behavioral traces of dislike or aversion (Grahek 2007). We will take a closer look at this condition in the next section. A fourth dissociation occurs in certain cases of injury on the battlefield or in sports competitions. Agents in these contexts may sustain even severe injuries while continuing in their endeavors undisturbed. When this happens, there are N-stimuli, N-events, and N-processes that are not fully registered in the cortex or the limbic system (Melzack et al. 1982).

When we consider these dissociations in relation to the folk theory of pain, we realize that we lack the conceptual resources to describe them efficiently and effectively. The folk theory provides us with a single relevant concept, the ordinary concept of pain, and it is clearly impossible to keep track of all the dissociable components of pain with that limited vocabulary. This suggests that we would be significantly better off, in philosophy and psychology, as well as in the clinic, if we were to replace that single concept with at least five others with more narrowly focused reference.

In addition to being conceptually impoverished, the folk theory of pain verges on incoherence in its handling of the question of whether pains are mental or bodily. Folk psychology authorizes us to insist with every possible degree of authority that we are actually in pain when we are suffering from phantom limb pain. It will not permit us to say that we are undergoing a hallucination of pain,

for any such claim would presuppose an appearance/reality distinction with respect to awareness of pain, and folk psychology does not countenance such a distinction. From the perspective of folk theory, if one is having an experience as of pain, then one definitely is in pain. It follows that pain is mental. After all, there are no relevant bodily phenomena in cases of phantom limb pain—there are no relevant N-events and no relevant N-processes. This pulls us in the direction of thinking that pains are N-states, or, at least, supported by N-states. But there is another side to this story. When we are aware of a pain, it always seems to us that we are aware of it as having a bodily location. The folk theory fully endorses this impression, encouraging us to use a variety of spatial concepts in characterizing pains. (We can speak of the size of the area of the body where pain is occurring, the spread of pain from one area of the body to another, the intensity gradient across an area, and so on.) To be sure, folk psychology allows us to recognize the existence of phantom limb pains, and thereby allows us to think of pains as occasionally having illusory locations; but it could not comfortably accommodate the proposition that the locations of pains are always illusory. The adjustments would be severe. This pulls us in the direction of thinking that pains are N-events. Moreover, it is a fact of great importance that pains are objects of awareness. We can attend to pains, pick them out with demonstrative concepts, and use those concepts to think about them. But if they were N-states, as opposed to N-events, our only awareness of them would have to be a kind of second order introspective awareness, and it is very implausible that we are engaging in introspection whenever we are aware of pains. It would, for example, follow that animals and small children cannot be aware of pains, for it is implausible that such creatures have the second order conceptual vocabulary that introspection requires. (For animals, see Cheyney and Seyfarth 2008.) In addition, it is implausible that adult human brains are constantly engaged in introspective monitoring. Introspecting places additional demands on energy resources and is not to be undertaken lightly.

Of course, if the foregoing claims about the flaws of the concept of pain are correct, any theory that both calls itself a theory of pain and is fully coherent, in the sense that it is not pulled in different directions by different forces, will inevitably pull the concept loose from some of its moorings. It follows that, in a sense, there cannot be a coherent theory of pain. I accept this conclusion, but I don't think it absolutely precludes the enterprise of determining what pain is—of finding out what it consists in. This is because there can be a theory of what pain is that *rationally reconstructs* the concept of pain—that is, a theory of what pain is that takes the form of a proposal about how to use the concept of pain in the future. It is this reconstructive enterprise that I will be pursuing in the present chapter.

To amplify: I will eventually claim that pains are identical with N-events. When I make this claim I should be interpreted as *recommending* that we so use "pain" in the future that "pains are identical with N-events" is true. The recommendation

will count as a theory of *pain* because it will capture several (though not all) of the strains in the use of the commonsense concept. And, I suggest, it will be a well motivated theory because the strains it will capture are of the first importance. Thus, it will capture the part of the concept that represents pains as having bodily locations, the part that represents awareness of pain as a first-order bodily awareness, and the part that represents pains as intimately linked with bodily damage. All things considered, having a theory that preserves the folk intuition that awareness of pain is not governed by an appearance/reality distinction is much less important, especially since, as we saw in Chapter 5, there is reason to insist on an appearance/reality distinction with respect to other forms of qualitative awareness.

The conclusion that the folk concept of pain is incoherent, being pulled in different directions by irreconcilable forces, is challenged in a recent article by Emma Borg and colleagues (Borg et al. 2020; see also Liu 2021). The authors recognize that our use of the concept has both a mental and a bodily dimension, but they suggest that the use is governed by weighting parameters that take on different values in different contexts: "in different contexts different elements of the concept could be activated, enhanced or suppressed" (p. 31). Since different facets of the concept are operative on different occasions, the concept can be perfectly coherent. It would only be incoherent if its mental and bodily dimensions were operative at the same time. This view is interesting, and it is nicely developed by the authors, but I doubt that it is correct. The deepest stresses in the use of the folk concept emerge only in special clinical settings, such as those involving phantom limb pain, and it seems quite unlikely that such cases played a role in shaping the concept. It is more plausible that the concept is designed to function in normal settings and that its different strands come apart when we are confronted by dissociations.

At all events, in the next section I will initially eschew the folk concept and focus on what is going on at the neural level when folk psychology authorizes us to speak of pain. Thus, initially, I will confine my attention to N-responses, N-states, N-processes, N-events, and N-stimuli. The goal will be to work out the relations among the parts of what I will call the *damage detection system*. A bit later on, proceeding in the spirit of rational reconstruction, I will make a proposal about how pain figures in this system. I will then defend this recommendation against a range of objections.

Before turning our backs on the folk theory of pain, we should note that one aspect of it is well founded and should be retained in any attempt to improve on the theory. Folk psychology is correct in representing pains as having bodily locations. This is sometimes denied. Thus, it is sometimes maintained that the "locations" of pains have no immediate connection to parts of the physical body. Rather, to say that a pain is "located in" one's left hand is just to say that the pain has a certain quale Q such that (i) it shares Q with certain other pains, certain

itches, certain burning sensations, and so on; and (ii) sensations with Q are distally caused by various processes that are physically located in one's left hand. Thus, according to this picture, awareness of sensations as having Q does not involve awareness of sensations as having a bodily location. It is awareness of a quality that stands in a contingent causal relation to events in a part of the body. As such, awareness of Q is comparable to seeing smoke that is caused by a fire that is not in one's line of sight. (Matthen 2021 is an exceptionally clear locus for such a view, though Matthen is concerned with sensations of touch rather than pains.)

This deflationary picture has a certain appeal, but it ignores three facts: first, the fact that the brain desperately needs information about the bodily locations of actual and potential damage; second, the fact that awareness of pains and their locations constitutively involves representations; and, third, the fact that the contents of these representations, like the contents of all representations, are determined by the informational needs of the organism. It follows from these facts that awareness of pains as located really is awareness of them as occurring in parts of the physical body. (I will revisit this line of thought in the next section.)

III. The Damage Detection System and Its Relation to Pain

N-states represent N-events. To appreciate this, note first that since N-states are caused by N-events, they encode information about them. Of course, they encode information about other occurrences as well—for example, activity in the spinal cord that is caused by N-events, and activity in the thalamus that is caused by the given activity in the spinal cord. But whether a representation refers to an item X is not determined only by informational relations. It also depends on the value that the information has for the organism; and in the present case, it is absolutely crucial that the organism have information about N-events. Information about activity in the spinal cord and the thalamus is useful only insofar as the activity mediates the flow of information from N-events to the insula and other relevant brain regions. Thus, on the one hand, considered independently of their relations to N-events, neither activity in the spinal cord nor activity in the thalamus is of any intrinsic significance to the organism. The same is true of activity in the non-initial segments of A-δ fibers and C-fibers. On the other hand, the brain desperately needs information about the magnitude and location of actual and potential damage to the body, so it can act to fend off harmful stimuli at the relevant locations and commit to nursing the afflicted areas. N-events encode information about these matters. It is plausible, then, that N-states have the function of encoding information about N-events. In combination with the dominant theme in current theories of representation (Millikan 1984, Dretske 1995, Neander 2017, Shea 2018), this conclusion implies that N-states represent N-events.

The case for this conclusion is strengthened by (i) the fact that the cortical apparatus that supports N-states incorporates somatotopic maps, (ii) the fact that locations on the maps are wired to peripheral and internal nociceptors in the corresponding bodily locations, and (iii) the fact that scientists have found it fruitful to think of the locations on the maps as representing activity in the peripheral regions of the body to which they are wired, as opposed to neighboring regions like the thalamus or comparatively nearby regions like parts of the spinal cord.

To fill out this portrait of the damage detection system, N-states are able to register the locations and damage-potential of N-stimuli by virtue of representing N-events, since N-events are caused by N-stimuli and to a large extent reflect their properties. The other main feature of N-states is that they cause N-responses, limbic activity that plays a large role in initiating behavioral responses to N-stimuli—that is, such responses as disengagement from N-stimuli and nursing of the aversively stimulated body part.

Where is pain located in the damage detection system? I suggest that we can answer this question by reasoning as follows:

1. In the damage detection system, N-states represent N-events.
2. States of awareness are constituted by representational states.
3. N-states are strongly correlated with states of awareness of pain, and are therefore the best candidates to be the representational states that constitute such states of awareness.
4. Therefore, by 1–3, states of awareness of pain represent N-events.
5. The object of a state of awareness is the item that the state represents.
6. Pains are the objects of states of awareness of pain.
7. Therefore, by 3, 4, and 5, pains are N-events.

We have just been witnessing a rationale for step 1 of this argument. Step 2 is supported by the representational theory of awareness, which is defended in Chapter 1 and Chapter 5. Step 3 is strongly confirmed by psychophysical studies (Price 2000, Craig 2010, Craig 2015). Step 5 is also authorized by the representational theory of awareness. And, finally, 6 is a truism. In view of all this, it is appropriate to conclude that the argument is sound. Moreover, it is not much of a stretch to move from its conclusion to the further claim that pain is identical with the property *being an N-event.* And that is the theory of pain that I wish to propose.

In Chapter 1 I advocated for representational pluralism, the view that a single state can stand in different representation relations to different properties. More particularly, a representation can represent a property by which it is causally entrained and at the same time represent a property that reflects the ecological value of the representation—that is, reflects Mother Nature's reasons for endowing us with the representation in the first place. Thus, it might be that activity in a

toad's T5-2 cells both represents the complex of properties *small*, *elongated*, and *moving along its longest axis*, and also represents the ecologically relevant property *food* (Neander 2017). Although I will not emphasize this in what follows, I think we should recognize a similar representational duality in the contents of N-states. The property that causally entrains N-states is *being an N-event*, and the ecologically relevant property is something like *harbinger of damage.*

IV. Six Concerns about This Theory of Pain

First, it might seem that the identification of *pain* with *being an N-event* is challenged by the fact that we have no inkling of nociceptive cells or activity in such cells on occasions when folk psychology authorizes us to claim awareness of pain. But this concern is addressed by a point made in Chapter 5—specifically, the point that the representations that figure in awareness of any kind always motivate an appearance/reality distinction. In this case as in the others we considered earlier, there is a distinction between the way an item is represented as being and the way the item is in itself. Because of this, an identification of the quale *pain* with *being an N-event* has no tendency to imply that the internal natures of N-events should be accessible to us when the quale is encountered.

Second, if *pain* is identical with *being an N-event*, then pains can exist even when agents are not aware of them. To see how this would happen, consider a subject whose spinal cord has been damaged in a way that prevents the upwards transmission of signals coming from nociceptors in the subject's right foot. In such a case, aversive stimulation could cause nociceptors in the foot to fire without triggering a cortical representation, and therefore without giving rise to awareness of the quale.

Our folk intuitions are challenged by the idea that there are pains that are not recognized, but reflection shows that no apology is called for. It is a central claim of the general approach to qualia recommended in Chapter 5 that awareness of qualia requires representations of qualia. Since representations are in general dissociable from the items they represent, it is a central feature of the approach that qualia can exist without being objects of awareness. Indeed, this is a feature of any theory which maintains that qualia are objects of awareness rather than intrinsic properties of states of awareness. Thus, the feature is shared by theories claiming that awareness of qualia involves being acquainted with them. The only theories with any following that lack the feature are versions of adverbialism, and we have found that such theories are very seriously flawed.

A third concern has to do with phantom limb pains. It seems initially plausible that we are aware of the quale pain in such cases, but how can this be if the quale is identical with activity in remote nociceptors? The right way to think about this issue, I suggest, is to recognize that the pressure to say that phantom limb pain

involves awareness of the pain quale comes entirely from the folk picture of pain, and, in particular, from the part of that picture which denies that awareness of the quale is governed by an appearance/reality distinction. On the alternative picture that I am recommending, it is perfectly appropriate to say that in phantom limb pain, there is an appearance of the quale without an accompanying reality. In short, on that alternative picture, phantom limb pains involve hallucinations. They are, to be sure, hallucinations with unfortunate consequences, for they trigger a storm of negative responses, such as aversion, but, as is well known, it is often the case that hallucinations cause emotional upheavals.

Fourth, there is a concern that has to do with the perceived intensities of damage signals (Pautz 2010). These perceptions are less well aligned with levels of activity at the tips of C-fibers and A-δ fibers than they are with levels of nociceptor-induced activity in the spinal cord, and they are less well aligned with levels of activity in the spinal cord than they are with levels of nociceptor-induced activity in the thalamus and elsewhere in the brain (Porro et al. 1998, Hill 2014b). Levels of nociceptor-induced activity are modulated by a range of factors as they proceed upwards to the cortex, and, as a result, they are often significantly higher or lower than the corresponding levels of activity in remote nociceptors. These facts might suggest that we should say that the relevant cortical representations stand for activity in the spinal cord, or in the thalamus, or perhaps in the cortex itself. After all, it can seem that perceived intensities of the represented items should match the levels of activity exhibited by those items.

Part of the answer to this is that we are required by a fundamental principle of representational content to say that N-states stand for N-events. According to the principle I have in mind, content must involve entities about which the representing mind/brain needs information. As we noted earlier, the brain desperately needs information about sites of bodily damage so that it can devise and execute plans for diminishing and coping with that damage. That is, if there is damage in region R of the body, the brain must have information indicating the presence of damage in R so that appropriate actions can be directed on R and objects impinging on R. The brain need not have information about activity in the spinal cord or the thalamus because those areas are not the sites at which action should be directed. (Nor are they really sites at which action *can* be directed.) It follows that the best candidates for the items represented by cortical representations are located in far-flung regions of the body, which means that the best candidates are N-events.

This is only part of a reply to the intensity concern, for we can still wonder why the perceived intensity of activity in remote nociceptors should be out of line with the actual intensity. There are several answers to this question. In some cases, the discrepancies can be understood as errors resulting from damage or disorder in nociceptors or in the higher-level neural structures that conduct nociceptive information. For example, neural structures in the spinal cord can become

sensitized to signals from nociceptors in a condition known as "wind up." This may cause them to magnify the signals coming from nociceptors, and it can also produce spontaneous pain, in which intense pain qualia are present even when peripheral nociceptors are not firing (Bushnell and Jessell 2013). Since in cases of this sort the damage detection system isn't functioning properly, it is natural to describe the divergence between perceived intensities and peripheral activity as misrepresentations.

In addition to cases of this kind, which involve special circumstances, there are also systematic divergences. Since they are systematic, it is less plausible that these divergences can be classified as misrepresentations. Misrepresentation should be the exception, not the rule.

To amplify, in a series of classic papers (e.g., Stevens 1962), S. S. Stevens argued that the perceived intensities of pains are related to the intensities of peripheral stimuli by power functions. Stevens further maintained that the exponents of these functions are in some cases significantly greater than 1, and that this is true, specifically, in the case of the perceived intensities of pains that are produced by electric shocks. Thus, according to Stevens, the exponent of the function describing the relationship between perceived intensities of pains and the corresponding electric shocks is 3.5. Now, there is room for doubt as to whether this exponent provides the optimal interpretation of the relevant data. Moreover, there is reason to think that we need much smaller exponents, and in fact exponents that are quite close to 1, to describe the relationship between the perceived intensities of other forms of pain (e.g., pains caused by heat) and the intensities of the corresponding causes (Adair et al. 1967, Craig et al. 2001). It must be acknowledged, however, at least as a possibility, that the perceived intensities of pain of certain types are never proportional to the intensities of the corresponding stimuli. In a phrase, it may be that the perceived intensities of certain pains are characterized by what is known as *response expansion*. If so, they pose an additional challenge to the view that pains are peripheral states. (For an exceptionally clear and detailed elaboration of this argument, see Pautz 2010.)

As the reader may have anticipated, my response to this objection is that it is possible to accommodate response expansion, and lay the objection to rest, by supposing that what N-states represent is not the actual levels of activity in peripheral damage detectors, but rather quantities that are functions of them, where the functions are computed by the cortical information processing mechanisms that are responsible for N-states. In short, the point about response expansion can be deflected by a hypothesis like the Thouless property hypothesis that entered the book back in Chapter 2.

This is also how I would like to answer concerns about perceived intensities arising from the fact that they are influenced significantly by various high-level psychological factors, including attention (Bushnell et al., 2004), emotion (Rainville, 2004), and stress (Rhudy and Meagher, 2000). To a large extent these

factors operate centrifugally, by sending facilitatory or inhibitory signals down pathways leading from the brain to various segments of the dorsal horn of the spinal cord. These descending signals combine with messages from the periphery to determine the strength and distribution of the signals that will then ascend via afferent pathways in the spinal cord to the brain, eventually producing an experience of pain. When one has an experience with an etiology of this sort, the perceived intensity of the pain is of course influenced by the relevant psychological factors, and is to that extent at variance with conditions on the surface of the organism. But these facts about perceived intensities are easily accommodated by the hypothesis that perceived intensities are the values of functions that are computed by processes that precede the conscious experience of pain.

I won't say more about this hypothesis in the present chapter, but it should be understood to be part of the package I am proposing.

Fifth, all tokens of the property pain present themselves as similar to the other tokens in at least one respect—they are all pains. That is, they seem to have a common nature. But the theory I am recommending equates the property *being a pain* with the disjunctive property *being a case of activity in the tips of C-fibers or in the tips of A-δ fibers or both.* How can pain seem unified if it is in fact disjunctive?

This concern would pose a serious problem if I were not presupposing a representationalist account of phenomenology and phenomenal awareness; but when seen from a representationalist perspective, the concern has an easy answer. As we have observed again and again, representationalism allows us to distinguish between the appearance of qualia and the corresponding reality. Recall that in Chapter 5 I was able to use this distinction to give a straightforward explanation of why the *quale* yellow seems simple and unanalyzable when, on any reasonable theory of the nature of the quale, it is actually extremely complex. The explanation had two parts: first, the proposition that the representation that supports awareness of the quale might be simple or atomic, relative to the system of representation of which it is a member; and, second, that if it is, it inevitably lacks the ability to register the various different components of the phenomenon that it represents. A closely related answer can be given here. The representation that supports awareness of the property *being a case of activity in the tips of C-fibers or in the tips of A-δ fibers or both* does not have to be disjunctive, or complex in any other way, in order to fulfill its representational function; and if it is in fact atomic, relative to the system of which it is a member, the states of awareness it supports must necessarily fail to provide access to the complexity of its representatum. (A similar answer can be given to the question of why awareness of visual and other perceptual metamers is blind to their differences.)

A final concern has to do with the choice of N-events to serve as the representata of the cortical representations we have been considering. There is an alternative choice—namely, N-stimuli, the states of damage and disorder that cause peripheral and internal nociceptors to be activated. Why not suppose that

N-stimuli are the principal objects of awareness when folk psychology authorizes us to claim awareness of pain?

We should prefer the hypothesis that pains are N-events over the hypothesis that they are N-stimuli for two reasons. First, N-stimuli include things like gas pressure, electric shocks, tumors, and secretions of stomach acid. It would be decidedly odd to classify phenomena of these kinds as pains. Second, pain presents itself as occurring within the body, either just inside the surface of the skin or at a deeper level of the interior. We can only access it via interoception. Activity at the tips of C-fibers and activity at the tips of A-δ fibers have an appropriate internality. On the other hand, we can often access actual and potential damage with externally oriented senses. We can often see damage and touch it with our fingertips.

What about the hypothesis that pains are N-processes—action potentials running from peripheral sites to the spinal cord? Is there any reason to prefer it to the hypothesis that pains are N-events? Not that I can see. Moreover, there is a decisive reason for rejecting it. Unlike N-processes, which extend along nerve fibers that in some cases are several feet long, N-events occur at the sites of damage. These sites are also the locations of pains. (I am indebted here to Jack Lyons.)

In view of these considerations we should conclude that N-states represent N-events, but this does not preclude us from also saying that N-states represent N-stimuli. Recall that in Chapter 1 we found reasons for endorsing representational pluralism, the view that representations can have two or more representata. Roughly, the idea there was that a representation R might have the function of encoding information about the property P that is causally responsible for its being tokened while also having the function of encoding information about a property Q that is linked probabilistically with P, provided that Q figures in an explanation of why R was selected. This idea is applicable here. It is clear that the system of cortical representations that figures in the damage detection system was selected because of its probabilistic relation to damaging N-stimuli no less than because of its causal relation to activity at the tips of C fibers and A-δ fibers. That is, it is clearly possible to give explanations for the selection of the system from two different perspectives, both fully legitimate. Accordingly, it seems appropriate to say both that N-states represent N-events and that they represent N-stimuli.

This completes my attempt to explain pain and awareness of pain in terms of the damage detection system involving N-responses, N-states, N-processes, N-events, and N-stimuli. The main idea is that what folk psychology describes as awareness of pain is a form of awareness that constitutively involves cortical representations of activity in peripheral and internal nociceptors. Because this form of awareness involves representations, there is an accompanying appearance/reality distinction: it can seem that one is aware of an N-event even when no N-event is occurring, as in cases of phantom limb pain, and an N-event can occur even though there is no awareness of it, as sometimes happens on the battlefield and in

athletic contests. Also, the appearance/reality distinction explains why we are not aware of the complex internal natures of N-events when we are aware of N-events as occurring, just as the corresponding distinction explains why we are not aware of the complexities of yellow light when we are experiencing a yellow phenomenology. On this account, the *quale* pain turns out to be the property *being an N-event*.

For expository reasons I have assumed throughout this section that the quale *pain* cannot be resolved into two metaphysically independent qualia, one having to do with the sensory nature of pain and another having to do with hurting. I turn now to the question of whether this assumption is correct.

V. Hurting

There are several conditions in which subjects describe themselves as experiencing pain but in which they appear not to be bothered by their experiences. This is true of subjects who have undergone lobotomies, subjects who have undergone cingulotomies (surgeries in which portions of the anterior cingulate cortex has been removed), and subjects who have been given large doses of morphine (Foltz and White 1962, Brand and Yancy 1997). It also appears to be true of experienced meditators. (I've had two meditation teachers who had no need for novocaine during dental surgery. For more on this topic, see Grant 2014.) The most striking cases of dissociation of pain from an aversive reaction involve the condition known as *pain asymbolia*, a disorder that can occur when there is damage to the insula (Basbaum 2021). I quote from a study of six asymbolia patients (Berthier et al. 1988, 43):

> Although all 6 patients could adequately recognize painful stimuli and distinguish sharp from dull, all of them showed a lack of *response* to painful stimuli.... Neither superficial nor deep pain stimulation elicited a motor withdrawal, grimacing, or an appropriate emotional response. One patient not only failed to show a withdrawal response but also exhibited a reaction of "approach" to the painful stimuli (i.e., he directed his limb toward the noxious stimuli). Inappropriate emotional responses were common: patients smiled or laughed during the pain testing procedure. This abnormal behavior ceased abruptly on discontinuing stimulation. All patients appeared quite unaware of their abnormal reactions and seemed unable to learn appropriate escape or avoidance responses. None of them became anxious or angry during the pain testing procedure; in fact, while all could recognize pain, none of them reported any unpleasant feeling. Patients showed normal autonomic reactions (tachycardia, hypertension, sweating, mydriasis (dilation of the pupil)) during the painful

> stimulation, but failed to react with a flinch, blink, or adequate emotional responses to threatening gestures presented to both hemispaces. Five patients also failed to react to verbal menaces.

The six patients in this study are representative: asymbolia patients typically insist on describing their experiences as pains and are able to discriminate sensations occasioned by aversive stimuli from sensations of other kinds, but they deny that they are bothered by their pains, and these denials are borne out by their nonverbal behavior. (For further discussion of these patients, see Schilder and Stengel 1931, Grahek 2007, 45–7, and Klein 2015, 2017.)

A number of authors have felt that the dissociation between pain and aversion in asymbolia patients indicates that the pain qualia we experience in normal cases actually have two aspects or components, a sensory component and a valenced component that I will call *hurting*. Cutter and Tye (2011, 92) make a case for this view in the following passage:

> Many commentators have noted . . . that if tissue damage is all that pain experience represents, then it's hard to explain one salient aspect of its phenomenal character, namely its "negative affect." Pains don't merely inform us of the presence of some disturbance at a location in our bodies; they hurt. A theory which says that pain experiences merely represent tissue damage at some location in one's body would seem to leave out the painfulness of pain. This is not to say that pain experiences do not represent the physiological condition—the physical disturbance—of the relevant body parts. But that cannot be all there is to such experiences. For the physiological disturbance that you feel in your arm when you experience pain in your arm also feels bad.

Aydede has a similar position (2019):

> Pains are not only sensory or perceptual experiences, they are also affective-motivational experiences, or at least they seem to have an affective aspect. Feeling pain is normally having an awful, hurtful, 'painful,' experience. So we may say that pains have a negative hedonic valence or affective value.

Other authors who endorse the view include Dennett (1978), Hardcastle (1997), Price (2000), Grahek (2007), Bain (2013, 2017), Klein (2015, 2017), and Kozuch (2020).

According to these authors, then, the ordinary experience of pain has two qualitative dimensions, a sensory dimension and an affective dimension. It is in virtue of the latter that pains hurt. Moreover, it is part of the view that the two dimensions are nomologically separable and in fact are separated in certain unusual circumstances.

I have for some years held an opposing view, according to which there is only one quale associated with pain—the sensory quale (Hill 2009, 2014a, 2014b, 2017). (To simplify the exposition, I will assume in this section that there is only one form of sensory pain—even though, as we have seen, sensory pain actually comes in a number of determinate varieties.) As I see the matter, it is just an illusion that pain asymbolia forces us to recognize two pain qualia.The illusion occurs because in asymbolia the sensory quale has lost many of its causal powers—specifically, its ability to cause the negative desires and emotions that accompany pain in normal cases. It has lost its ability to *bother* subjects. In other words, as I see it, instead of there being two pain qualia, a sensory quale and a hurting quale, there is only one, a sensory quale, and there is duality only because the sensory quale has two quite different sets of causal powers, displayed respectively in normal subjects and in asymbolia patients.

I have two reasons for favoring this deflationary view. One is that it seems to be the view that is most in keeping with the testimony of asymbolia patients. They say that they continue to feel pains but that their pains no longer *bother* them (Grahek 2006). More importantly, though, the dual qualia view is incompatible with the fact hurting is *intrinsically motivating*—that is, the fact that insofar as a pain hurts, one has immediate, visceral, and galvanizing motivation for trying to end it, or to at least reduce it. I mean to be making a very strong claim here, namely, the claim that it's *impossible* for a subject to experience a pain as hurting without being motivated to do something about it. (Of course, if the pain lacks intensity, the motivation is correspondingly weak. Also, what one does about hurting may be no more than lying down and shutting one's eyes, as victims of migraines are wont to do.) Now hurting would not be intrinsically motivating if it was a "factual" property like *harbinger of damage* (Cutter and Tye 2011) or *harbinger of a reduction in reproductive fitness* (Kozuch 2020). An asymbolic might be aware that their skin was being damaged and be indifferent to the fact, or even mildly amused. And someone might notice a reduction in reproductive fitness and be delighted. (Consider the vasectomy patient who shouts with joy, "Relief at last! No more worries about unwanted children!" (I am indebted here to Leonard Katz.)) More generally, factual properties of objects do not dictate our attitudes toward objects, and this is no less true of factual properties of pain than of factual properties of external objects. On the contrary, motivation comes from our *attitudes towards* facts. It stems from our evaluations, and our positive and negative attitudes. It isn't a matter of registering the existence of objects and their factual or "natural" properties.

It follows from this that if we want to determine what hurting consists in, we should consider the attitudes that are directed on pain. In normal cases there will be a number of these attitudes, but it seems that the most natural candidate is the attitude of *dislike*. It is clear that dislike is intrinsically motivating: it holds as a

general rule that to dislike something is be motivated to disengage from it and to avoid it in the future. Indeed, if dislike is strong enough, as it normally is in cases of pain, it is experienced as a *demand* that one take steps to quell or withdraw from the object of dislike.

The crucial thing to note here is that if hurting consists in an attitude that is directed on pain, there is no need to suppose that there is a separate quality of hurting, a factual property that accompanies sensory pain in normal cases. The target of the attitude can be sensory pain itself.

Alas, there is no such thing as a free lunch. Like all philosophical positions, the present view has implications that can be questioned. The most significant of these is the central claim—that hurting is constituted by our dislike of pains. This claim is opposed by an intuition to the effect that the dependency between hurting and dislike is actually causal rather than constitutive. Moreover, according to the intuition, dislike depends on hurting rather than the other way around. We dislike pains because they hurt.

To respond to this objection, there is reason to doubt that folk wisdom is a trustworthy guide to dependency relations among pains and affective responses. If the relation was causal, there would have to be at least a minute temporal interval between the onset of hurting and the onset of dislike. But introspection is too blunt an instrument to detect such a gap. Our sense that hurting causes the affective response to pain is not due to introspective measurements, but rather to the observation that information about pain travels from the body to the mind, and to a vague impression that hurting is somehow bound up with sensory pain. This impression is of course correct, but, when we reflect on it, we appreciate that it does not imply that hurting is reflected in the flow of information from the periphery. Still less does it imply that hurting is an independent factual property of sensory pains. It is fully compatible with the claim that hurting is an attitude that is brought into being by sensory pain and is directed on it.

Another objection to the present proposal invokes an alleged intuition to the effect that hurting is an object of awareness—a property of sensory pains that is given to us when we are aware of them. If hurting is an object of awareness that we take in when we are aware of sensory pains, the objection maintains, it must be a property of sensory pains, not an attitude toward them. My response is that this intuition should be ignored in constructing a theory of pain, for it is defeated by other considerations. If we are aware of hurting as a property of sensory pains in everyday experience, then, it seems, asymbolics should have difficulty in recognizing pains as such after the onset of their asymbolia. But their testimony reflects no such problem. On the contrary, as we have seen, they assert that their pains are qualitatively the same after they have ceased to hurt. Moreover, the intuition is challenged by an opposing intuition that seems much more compelling—namely, the intuition that hurting is intrinsically motivating. Awareness involves

representations, and it is pretty clear that a naturalistic account of representation requires that represented properties be natural or factual properties. Factual properties are not intrinsically motivating.

There have been attempts to account for the intrinsically motivating character of hurting while honoring the idea that hurting is an independent object of awareness. To illustrate, e*valuationism* asserts that hurting is the normative property *bad*. It claims both (i) that badness is an object of awareness and (ii) that awareness of pains as bad is intrinsically motivating (Bain 2013, 2017). But there is a problem with this view. Given a representational theory of awareness and a naturalistic account of representation, claim (i) presupposes that badness is a natural or factual property. It is a substantial cost of evaluationism that it presupposes this highly controversial proposition. Another example is the version of *imperativism* which asserts that hurting is an imperatival property such as *must be stopped*. It is plausible that awareness of such a property would be intrinsically motivating, but I see a problem with the idea that imperatival properties are sufficiently natural or factual to serve as representata. Indeed, one could be forgiven for doubting that imperatival properties exist. (There are discussions of imperativism in Klein 2015 and Klein 2017. The version of imperativism under consideration here is suggested by some parts of those discussions.)

A reader might worry that the present account of hurting leaves us without a way of accounting for masochism. Masochists pursue pain. How can this be true if hurting is dislike and dislike provides a motive for disengaging from painful stimuli? I think we can see the answer if we reflect on examples of other kinds. Consider a situation in which a billionaire offers a million dollars to anyone who will perform a universally detested action, such as cleaning out a chicken coop. Clearly most of us would be tempted to accept the offer. In this case as in many others, including no doubt cases of masochism, dislike is overridden by countervailing needs or desires.

The idea that hurting consists in disliking sensory pains originated with Kent Berridge and his associates, though I hope I have increased the clarity of the view, and have also added to the reasons for embracing it (Kringelbach 2005, Smith et al. 2010, Kringelbach and Berridge 2015, Kringelbach and Berridge 2017, Nguyen et al. 2021). I should also mention that there are writers who propose attitudinal theories of hurting that diverge from the one recommended here. The reader is referred to Heathwood 2018 and Jacobson 2019 for interesting alternative versions of the underlying idea.

Before concluding this discussion of hurting, I would like to say something about an intuition that is quite natural but is nonetheless deeply misleading. This is the perception that hurting is partially constitutive of pain. According to this perception, hurting and pain are metaphysically inseparable. The perception can seem compelling: When one is experiencing a pain, the thought that what one is experiencing could be broken into two components would seem absurd. That is,

when one is experiencing pain, it seems to be part of the very nature of what one is experiencing that it hurts. Moved by this intuition, one might easily be led to deny that asymbolia patients really feel pain. Rather, one might think, they experience a different sensation that is similar to pain in some respects. Their tendency to classify this other sensation as pain is due in part to this similarity and in part to its having the same causes as pain. Instead of speaking of pains, they should instead be using the expression "partial pains" or the expression "quasi-pains."

It is difficult to interpret the testimony of asymbolia patients since their experience is so different than the experience of most of us. Not having been in their shoes, it is easy for us to discount their testimony. It is therefore helpful to consider cases that are closer to home in which sensations take on different aspects. There are plenty of cases that meet this description: it frequently happens that a sensation undergoes a change in aspect and yet we can recognize it as the same. Perhaps the most salient example is the taste of beer, which seems pretty awful when first encountered but can come to be a favorite. Other examples include the taste of anchovies and the taste of Roquefort cheese. These examples involve transitions from disliking to liking. But changes in evaluation of tastes can also go in the other direction, as is best illustrated by the well-studied phenomenon of conditioned taste aversion. (A conditioned taste aversion is an enduring dislike of the taste of a food that one was eating just before falling ill, most typically before falling ill of a disorder that causes nausea. There need be no causal relation between the food and the illness for the aversion to be acquired, nor even the appearance of a causal relation. See, e.g., Garcia et al. 1955, Garcia and Koelling 1966, Seligman and Hager 1972, Bernstein 1978.)

To give an example of a different kind, for a number of years I disliked Cézanne's palette, much preferring Renoir's bright and highly saturated colors. But a radical switch occurred at some time I can no longer remember, and for a reason that I perhaps never knew. I now prefer Cezanne's palette. Similarly, many people dislike atonal music when they first encounter it, but come to appreciate it as their exposure to it increases.

In an ingenious and seductive essay (Dennett 1988), Daniel Dennett questions the evidence for preference reversals involving taste, maintaining that when we describe a new preference for a taste as a reversal, it could be true instead that the taste had actually changed and that we are being misled by a false memory. For example, according to Dennett, when we think that our dislike of the taste of beer has given way to a new attitude toward that taste, an attitude of approbation, we could be quite wrong. Instead of a preference reversal regarding a constant taste, it could be that a change in our taste buds has occurred, or a change in our olfactory bulb, with the result that the taste of beer is actually different. Our impression that the taste is the same could be due to a failure of memory.

This line of thought seems quite misguided to me. The problem is that there is no evidence that gustatory and olfactory qualia covary with changes in memory

in the way that Dennett's example requires. As the reader may know from tragic personal experience, or from newspaper articles about the consequences of Covid-19, the effects of Covid often include *parosmia*, a condition in which the olfactory qualia associated with odorants undergo profound changes. What is relevant here is that memories of former qualia are not affected by these qualitative changes. On the contrary, one of the main reasons why parosmia is tragic is that its victims are vividly and painfully aware that they have been deprived of qualia that were precious to them. (Cf. the testimony of parosmia victims at the BBC News site https://www.bbc.com/news/stories-55936729.) In view of this, we should see Dennett's suggestion that cases of preference reversal might be indistinguishable from cases of qualitative change as collapsing into a form of skepticism about memory. The suggestion presupposes that there might be a perfect correlation between changes in qualia and failures of memory. But the evidence speaks—indeed, shouts—against that idea. Accordingly, the suggestion can't be seen as a serious empirical hypothesis about actual mental processes. It has to be understood as an unsupported skeptical hypothesis about the trustworthiness of memory.

To summarize, when you're down on a type of taste, it can seem impossible that you could ever like *that taste*, and when you're an aficionado, it can seem impossible that you could dislike *that* taste. Equally, it can seem impossible that you could ever like *those colors* or *those chords*. But clinical and experimental data show that such reversals are in fact possible, and by the same token, they show that we are able to recognize sensations as the same despite having undergone a significant change in our attitude toward them. It is plausible that this is what happens in the case of pain asymbolia. The hurting dimension of pain can be pried off from it without changing the sensory dimension, or affecting our ability to recognize it as such. Moreover, just as it is legitimate for us to continue to refer to the taste of beer after it has undergone a change in aspect, so also is it appropriate for asymbolia patients to continue to refer to their sensations as pains. What has changed is their reactions to those sensations. The sensations themselves are the same.

VI. Awareness of Pain and Perceptual Awareness

I earlier floated the idea that awareness of N-events is a form of perceptual awareness. That may not be fully accurate. There are significant differences between awareness of N-events and the paradigmatic perceptual modalities. There are also important similarities, however, and they seem to warrant the thought that awareness of N-events is at least quasi-perceptual. I will list several of these common features (Hill 2014a).

First, all of the familiar forms of perception are associated with automatic attention mechanisms, and also with attention mechanisms that are under voluntary control. Among other things, as Marisa Carrasco has shown, such mechanisms can increase the resolution of our experience of an object of awareness, and heighten the contrast between an object of awareness and its background (Yeshuran and Carrasco 1999, Carrasco et al. 2000). The same is true of awareness of pain. We can attend to N-events, and, when we do, there is a higher level of resolution and also a more salient contrast between figure and ground.

Second, it is of the essence of perceptual representation to assign locations and other spatial characteristics to its objects. The same is true of awareness of N-events.

Third, there are a priori norms of good grouping that determine the ways in which perceptual elements are organized into wholes. For example, we group visually presented dots together if they are alike in some respect—that is, if they share a neighborhood in space, or a shape, or a color, or a size, or a common fate. The same is true of groups of N-events. Suppose that three N-events are occurring, two in your palm and one on your wrist. The two in your palm will seem to form a unified whole of a certain kind. Equally, two N-events that are alike in intensity, or that begin to exist at the same time, will seem to be members of a single "society," even if they lie at some distance from one another spatially.

Fourth, perceptual awareness tends to represent its objects as having highly determinate forms of the properties that it attributes to them. I can *believe* that an object is blue without believing that it is any determinate shade of blue. But I cannot visually perceive an object to be blue without perceiving it to be navy blue or some other highly determinate shade of blue. Or at least, this is true of the blue objects that I am foveating. There is another parallel here with awareness of N-events: I cannot experience an N-event as intense without experiencing it as having a particular level of intensity.

Fifth, perceptual awareness is particularized. I can form a *belief* that there are three books in a box without being able to uniquely specify any of the particular books that make my belief true. Equally, I can *believe* that someone has been eating my porridge without having any relevant beliefs that single out a specific individual. But perceptual awareness is different. If I am perceptually aware of the presence of three books in a box, I must be perceptually aware of properties (including egocentric properties) that uniquely determine each of the individual books. Equally, if I am aware of the existence of a trio of N-events in my arm, I must be aware of properties that uniquely determine each individual member of the trio.

Sixth, perceptual awareness has a certain mereological determinacy. I can form a *belief* about an object without forming any belief about its parts. But unless an object of perception is atomic, a minimum sensibile of some sort, I am inevitably aware of a range of its parts in being perceptually aware of the object, and also of

certain of the structural relationships among the parts. (I mean to be using "part" quite broadly here, so that it applies to temporal constituents of events and qualitative constituents of complex properties as well as to spatial parts of physical substances.) This is also true of awareness of N-events. N-events are normally experienced as extended in bodily space, and when they are so experienced, the parts are experienced as well, as are a number of the structural relationships among the parts.

VII. Conclusion

The commonsense concept of pain is used to keep track of several items that can come apart in laboratory and clinical settings. Accordingly, when we try to explain the metaphysical nature of pain, we are pulled in different directions. I have maintained that the strongest forces pull in the direction of identifying pains with activity at the tips of C-fibers and A-δ fibers, and have accordingly proposed that if we continue to speak of pains, our use of the term should respect that identity. I have also argued that the identity can be defended against a broad array of objections. In addition to this discussion of the sensory dimension of pain, there has been a discussion of its hurting dimension. The main conclusion of the latter discussion was that the hurting dimension is a certain kind of attitude towards a sensory state. Asymbolia patients differ from the rest of us in that this state is not activated when sensory pain occurs, and also in that they lack the emotions (anger, anxiety, etc.) that often accompany the state. Finally, I have maintained that awareness of pain shares several important characteristics with the paradigmatic forms of perceptual awareness. These common features suggest that awareness of pain should be thought of as quasi-perceptual in character.

7
Perceptual Consciousness

I. Introduction

We speak of perceptual experiences as having qualitative or phenomenal character, and we also speak of them as being conscious. In Chapter 5 I proposed an account of perceptual phenomenology. In this chapter I will continue to be concerned with perceptual phenomenology, but the principal focus will be on perceptual consciousness. What does it consist in?

I will be concerned in this introductory section with two types of consciousness that are widely acknowledged in the literature—*introspective consciousness* and *access consciousness*. They are of interest here because they can both be possessed by perceptual experiences. After describing them in the present section, I will proceed in later sections to consider their relationship to perceptual phenomenology, eventually maintaining that although phenomenology is closely connected with these two forms of consciousness, it is metaphysically independent of them. I will then turn to consider the view that perceptual experiences possess another form of consciousness—*phenomenal consciousness*—that is independent of introspection and access. This view is widely accepted, and I am among its adherents, though my take on phenomenal consciousness stands apart from the mainstream. My goal in considering this form of consciousness will be to identify its metaphysical nature—that is, to explain what it ultimately consists in.

Introspective consciousness is the property of being an object of introspective awareness. It is therefore like the property of being seen—a property that an item possesses just in case it stands in an awareness-relation to a subject. It is the property that Freud seems to have had in mind in distinguishing between desires (or wishes or beliefs) that are conscious and desires (or wishes or beliefs) that are subliminal or repressed. As he saw it, the former are mental states that the subject is aware of, while the latter are mental states that lie outside the scope of awareness (Freud 1965). Contemporary therapists have inherited this usage.

What is introspection? For present purposes, it suffices to say that it involves mental states that represent other mental states—that is, second-order representations—and also a range of other things, including working memory, attention, and contextually defined parameters that determine whether tokening a second-order representation is task relevant. I will say more about the nature of introspection in Section IX.

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0007

When the expression "access consciousness" was introduced some years ago by Ned Block, he defined it as follows: a representation is access-conscious if it is poised for use as a premise in reasoning, poised for rational control of action, and poised for rational control of speech (Block 1995, 231). In more recent writing Block has liberalized the definition considerably, saying that a representation is access-conscious if it is available for use by a broad range of higher cognitive agencies, including the ones that are responsible for reasoning, problem solving, planning, producing verbal reports, creating perceptual beliefs, and forming episodic memories (Block 2022). Roughly, a representation is access-conscious if it is "globally available" for use by a variety of high-level cognitive agencies. Similar ideas have been promoted by Baars (1997) and Dehaene (2014), and have found favor with many writers on consciousness, both in science and in philosophy. Evidently, there are widespread and vivid intuitions to the effect that there is a form consciousness that is constituted by relations to high-level cognitive states.

It is useful to distinguish between occurrent and dispositional forms of the two forms of consciousness we have been considering. Here are some definitions: (i) Occurrent introspective consciousness is the property that a mental state has when it is actually being registered by the introspective faculty of the relevant agent. (ii) Dispositional introspective consciousness is the property of being available to the relevant introspective faculty. (iii) Occurrent access consciousness is the property of actually causing one or more high-level cognitive states in the relevant agent. And (iv) dispositional access consciousness is the property of being poised for causing such states. I will assume that it is appropriate to think of all of these properties as forms of consciousness. Certainly they have all been cited as such in the literature (Block 2002, Gennaro 2012).

It is clear that all of these properties can be exemplified by perceptual experiences. In view of this, we can see that there are at least four different answers to the question with which we started, "What does perceptual consciousness consist in?" Moreover, each of these answers has merit. Can we conclude, then, that our inquiry has been brought to completion?

This is one of the questions to be addressed in the present chapter. I will consider several ideas about additional forms of perceptual consciousness and offer evaluations of them. Before we consider that topic, however, we should consider the question of how the four forms on our present list are related to phenomenology. A number of writers hold that one or more of these forms is partially constitutive of phenomenology: that every phenomenal state necessarily stands in actual or potential causal relations to behavior and/or downstream cognitive states, including introspective states. This view is widely held, both by scientists and by philosophers (e.g., Tye 2002, Dennett 2005, Cohen and Dennett 2011, Dehaene and Changeux 2011, Dehaene 2014, and Ward and Scholl 2015), but it seems quite wrong to me. I will maintain in the next few sections that

phenomenology is metaphysically autonomous—that it is in principle possible for it to exist without being accompanied by any of the relational forms of consciousness we have just been considering. Indeed, I will argue for a proposition that is much stronger than this one. As I see it, there are good grounds for thinking that experimental work has actually managed to separate phenomenology from actual and potential relations to downstream phenomena.

After arguing that phenomenology is independent of introspection and access, I will consider the possibility that there is an additional form of perceptual consciousness that warrants the name *phenomenal consciousness* because it is more intimately linked to phenomenology than either introspective consciousness or access consciousness. I will argue that we do need to recognize a form of consciousness that answers to this description. More specifically, on the account I will defend, phenomenal consciousness is the intrinsic, categorical base for introspection and access, which are both relational in nature. As we will see, this view comes to the same thing as saying that phenomenal consciousness is realized by very high levels of activity in certain areas of the brain, such as the fusiform face area, which is known to be intimately associated with experiential awareness of faces. It is part of this picture that we may have to recognize phenomenology that is not phenomenally conscious. Unconscious phenomenology would be realized by lower levels of activity in those same regions of the brain.

II. The Autonomy Thesis

The autonomy thesis is the proposition that phenomenology does not entail actual or potential relationships to downstream cognitive states. As noted, this proposition is highly controversial, with quite a few opponents ranged against it. It seems best, therefore, to assemble several arguments on its behalf. None of the arguments I will give is decisive by itself, but collectively they make a strong case, showing that the autonomy thesis can be defended from a variety of vantage points. I will give two arguments grounded in folk psychology in the present section, one that is based on introspective data, and another that is based on the role that phenomenology plays in the explanation of behavior. These two arguments are important because they explain the intuitive plausibility of the autonomy thesis. In the next section I will give a metaphysical, quasi-Leibnizian argument for the thesis, and in the section following that I will add an argument that stems from an important discovery in vision science. Hopefully every opponent of the autonomy thesis will find at least one argument in this set of four to have some lasting appeal.

Introspection provides us with our principal access to phenomenology, so it is natural to suppose that we can grasp some of the main properties of phenomenology by consulting intuitions about it that are grounded in introspection.

One of these intuitions is that perceptual phenomenology doesn't entail relations to concurrent cognitive states. Consider a case in which you are viewing a red object, and are therefore experiencing a reddish phenomenology. Suppose that you are introspectively aware of this phenomenology. Now it may be that you are also aware of a perceptual judgment to the effect that you are confronted by a red object, or a judgment that something looks red to you. You may also be aware of a decision to choose the red object to serve some purpose. It is clear, however, that these cognitive states are not in any way part of the phenomenology of your experience of red, though they may have their own proprietary phenomenologies. From the perspective of introspection, the phenomenology associated with your experience of the red object is exhausted by the appearance of red that you are enjoying. Relations to cognitive states have nothing to do with it.

This first intuition counts in favor of the independence of phenomenology from the occurrent forms of introspective and access consciousness. Another abiding intuition counts against the dispositional forms. According to this second intuition, phenomenology is *given* in introspection. We are *directly* aware of it. This calls into question the idea that phenomenology is partially constituted by dispositional properties. We are never directly aware of dispositions, though we may be directly aware of the non-modal properties that constitute their categorical bases, and are often aware of the concrete events that manifest them. We know about dispositions by inference from non-dispositional facts of these two kinds. But it is pretty clear that we do not have to resort to inference to apprehend phenomenology. The intuition of givenness is quite robust.

This brings me to the second reason for accepting the autonomy thesis, which focuses on the role that phenomenology plays in explanations of behavior. The key idea is that we sometimes need to appeal to phenomenology in explaining behavior even though the phenomenal states to which we appeal are not access conscious. The argument I will give is based on an example that originated with David Armstrong (Armstrong 1980, 59).

Consider a truck driver who "wakes up" to the realization that he has absolutely no memory of anything that has happened for the last 40 minutes. He has covered 30 miles of complex terrain, but his mind was fully occupied by other things—let us suppose that he was planning a vacation. Clearly, he must have been perceptually aware of the colors of traffic lights, the shapes, sizes, and distances of other vehicles, the curves in the road, and his own velocity and acceleration, for otherwise he would have collided with another vehicle or run off the highway. But, apparently, from beginning to the end, his states of awareness were not access conscious. Certainly they left no traces of themselves in memory. Indeed, the states of awareness really *could not* have been access conscious, since the driver's working memory was entirely occupied by plans for his vacation. Now in all normal cases, perceptual awareness that guides such complex behavior involves phenomenology. Accordingly, if we are to give an explanation of the driver's successful

navigation of the sort that we would give in other contexts, it is necessary to appeal to phenomenal states. And it seems reasonable to give an explanation of the same sort, because the behavior to be explained is of the same type.

As the reader may be aware from personal experience, unconscious behavior of this complex sort may occur even in contexts that are completely unfamiliar. Not all cases can be explained by appeal to learned dispositions that operate independently of high-level perception.

In evaluating this argument, keep in mind the fact that when drivers remember sequences of guiding perceptual states, they remember them as having proprietary phenomenologies. It seems that all that is missing in the case of Armstrong's driver is the memories. There is no reason to think that he lacked the kind of states that the memories are *of*.

III. A Metaphysical Argument for the Autonomy Thesis

Another argument for the autonomy thesis is suggested by the fact that there can be various levels of activity in the neural areas that are believed to support phenomenology. It is a natural thought that all of these levels can support some form of phenomenology. The phenomenology supported by higher levels of activity in these areas would no doubt be more robust than the phenomenology supported by lower levels, but diminishing degrees of robustness are compatible with a common kind affiliation.

There are a number of different ways of interpreting the notion of a level of neural activity. For example, levels might be rates of firing, but they might also correspond to numbers of participating neurons. Or they might involve a weighted average of these two quantities. Another possibility is that levels might correspond to the degrees to which patterns of firing are stable, where stability is understood to entail an ability to resist perturbation or disruption by other, competing patterns. A more complicated hypothesis is that there is a difference between levels of activity in regions of the brain that support feature detectors and levels of activity in nerve fibers that connect such regions, thereby supporting representations that attribute multiple features to objects. The need for such a distinction is suggested by studies of the phenomenon known as *crowding*, in which subjects are aware of features without being able to assign those features to determinate objects. On this hypothesis, as on related ones, the total phenomenology associated with a state would be a function of levels of activity of different kinds. There are still other ways of interpreting the notion of a level of neural activity. I will not favor any one of these interpretations here. The line of thought I wish to pursue prescinds away from these more concrete proposals.

That there can be different levels of activity in areas known to support phenomenology is illustrated by a famous experiment by Moutoussis and Zeki

(2002). Their study involved stimuli that had contrasting colors but were otherwise identical. When pairs of such stimuli are shown, respectively, to the right and left eyes, they are invisible, even though each would have been consciously perceived if it had been presented alone. Using fMRI, Moutoussis and Zeki were able to show that invisible stimuli are able to give rise to low levels of activity in visual areas that are known to support phenomenology when they are highly active (specifically, the fusiform face area and the parahippocampal place area). Moreover, this is just the tip of an iceberg. There are a number of other studies that concern unconscious activity in areas known to support phenomenology, including studies of binocular rivalry, backward masking, and flash suppression. (The reader is referred to Dehaene 2014 for a review.)

The fact that there can be different levels of activity in areas relevant to phenomenology suggests that phenomenology may extend to lower levels, albeit in a form that is less robust than the form that exists at higher levels. And, by the same token, it suggests that there is phenomenology that is not associated with either introspective consciousness or access consciousness. This is because subjects in studies like the one Moutoussis and Zeki conducted show no evidence of either of these forms of consciousness when low levels of activity in phenomenology-supporting areas of the brain occur. Such subjects cannot describe or recall the stimuli that cause the activity. In general, lower levels of activity in those areas seem to lack higher-level cognitive effects.

I said that the experiments *suggest* that different level of activity in the given brain areas may be accompanied by phenomenology that has different degrees of robustness. Is there any reason to think that this idea may be correct? Yes. The idea is supported by the following *continuity argument*, which may remind the reader of some things that Leibniz says about *petites perceptions* (Leibniz 1996):

First premise: The pains that fall within the scope of introspection are ordered by a relation of comparative intensity. Some pains are extremely intense, some are extremely faint, and there is a range of intensities that lie between these extremes.

Second premise: There is a lower bound on the intensities of pains that are introspectable.

Third premise: But this is only an epistemic lower bound. There is no reason to think that it is also a metaphysical lower bound.

Fourth premise: In general, when a natural magnitude comes in degrees, there are low degrees of the magnitude that fall beneath the scope of human awareness. Thus, for example, there are volumes of sound that are too low for the human ear to discern, and textures that are too fine grained for haptic exploration to detect—they seem perfectly smooth, but actually the surfaces have tiny grooves. Epistemic lower bounds on magnitudes do not normally correspond to metaphysical lower bounds. Normally, the metaphysical lower bounds extend below the epistemic bounds, and often far below.

Conclusion: All of this being true, there is reason to believe that the ordering of pains in terms of intensity extends beneath the reach of introspective discernment.

Similar arguments can be constructed for other types of phenomenology. Thus, perceived lights, perceived sounds, perceived pressures, perceived aromas, and perceived tastes are all ordered by intensity relations. In all cases there are epistemic lower bounds to these intensity orderings, but there is no reason to think that these perceptual phenomenologies are exceptions to the general rule that the epistemic lower bounds of magnitudes do not track metaphysical lower bounds.

The continuity argument falls considerably short of being a bullet-proof demonstration. It is based on an extrapolative inference from the many domains in which magnitudes tail off gradually, without regard for human epistemic capacities. Pain and other types of phenomenology could be exceptions. It seems fair to say, however, that it at least raises the probability of its conclusion. As a result of the argument, it is more likely that phenomenal states are ordered by a relation of comparative intensity or robustness that is in lock step with the ordering of levels of neural activity in brain areas that are known to support phenomenology. This in turn makes it more likely that phenomenology can exist independently of introspection and access.

At all events, while I will continue to assume in the background that phenomenology comes in degrees of robustness, and that the lower degrees fall beneath the thresholds of introspection and access, I will be focusing in most of the rest of this discussion on phenomenology that is highly robust. When I speak of phenomenology in the next section and beyond, unless otherwise indicated, I will always mean to refer to phenomenology that is at the top of the scale, even if I don't explicitly say as much.

IV. A Fourth Argument for Autonomy

I conclude this case for the autonomy thesis by describing and defending the "overflow argument" that has been developed by Victor Lamme and Ned Block. I will describe the argument briefly, and then dwell for a while on the question of what exactly it shows. The reader is referred to Landman et al. 2003, Block 2011, and Block 2022 for detailed expositions.

The overflow argument is based on a classic experiment by Sperling (1960) in which subjects were shown a matrix of twelve letters at an initial time T1. When interrogated at a later time T2, after the letters had disappeared, the subjects were able to name four of the letters but no more. They could, however, recall having seen a full matrix of letters. That is, they could recall having seen letters at T1 that they could not name at T2. By using a special technique that I'll describe in a

minute, Sperling was able to find experimental confirmation for the subject's claim to have seen all of the letters. Now the ability to articulate perceptual states by naming stimuli is a hallmark of access consciousness. Accordingly, Block concludes that the letters that were seen at T1 but not specifically recalled at T2 lacked access consciousness. If, as is natural, we also suppose that the initial (T1) awareness of all of the letters had a proprietary phenomenology, we can infer that there are phenomenal states that are not access conscious.

Sperling confirmed the testimony of his subjects by a cueing procedure. The letters in the matrix were arranged in three rows. Sperling sounded tones after the letters had disappeared from the screen. A low tone meant that subjects were supposed to name the letters in the bottom row; a medium tone meant that they were supposed to name the letters in the middle row; and a high tone meant that they were to name the letters in the top row. Although the subjects were only able to name four letters, it turned out that they were able to name the letters in *any* row that was cued by a tone, provided that the tone followed the initial display by a short enough interval (around 300 milliseconds). Evidently, the subjects must have *seen* all of the letters in order to be able to name the ones in an arbitrary row, and they must have *remembered* all of the letters at least to the point when the tone sounded. We must accept these hypotheses in order to best explain the subject's success in naming the letters in the row that had been highlighted by a tone.

I believe that Block's line of thought is ultimately successful, but it is easy to get confused in evaluating it. It presupposes a criterion for access consciousness that is arguably too demanding. According to the criterion, a phenomenal state fails to count as fully access conscious unless one can recall and name the content of every one of its parts. But one might think it no less reasonable to say that a state can count as access conscious if one can recall and name the *gist* of its content. This is relevant because Sperling's subjects could in fact recall the gist of their original experience at T1. They remembered that they had seen a matrix of letters. Hence, according to a remembered gist criterion of access consciousness, their initial experiences of the matrix were access conscious. Moreover, assuming that their total experiences could be factored into "smaller" experiences of individual letters, we can say that these smaller experiences were also access conscious. This is because the *gist* of each of the smaller experiences was remembered. The subjects remembered the unnamed letters *as* letters. (For discussions of gist, see Oliva 2005, Oliva and Torralba 2006, Haberman et al. 2015, Cohen et al. 2016, and Ward et al. 2016.)

So there are two possible criteria for access consciousness, one that is based on total recall of an earlier representation and a less demanding one that is based on recall of gist. Are these criteria rivals, and if so, is there a standoff between their respective proponents? Are the criteria equally well motivated? If so, Block hasn't succeeded in prying phenomenology loose from access consciousness.

The conception of access consciousness that is based on recall of gist can be tempting, but I think that in the end there is a strong reason to prefer the conception that is based on total recall. The reason for having a concept of access consciousness in the first place is that we want to be able to distinguish between those aspects of the representational content of a state that are available for use by higher-level faculties and those aspects that aren't. Let R be a perceptual representation, let C be the content of R, and let G be the gist of C. It follows that we should not say that R is access conscious just because G is available for use by higher cognitive faculties. That might suffice for the claim that R is *weakly* access conscious, but for R to be *strongly* or *fully* access conscious, all of C must be available for such use. To put the point another way, for R to be fully access conscious, all of the details of C must available to working memory. It isn't enough for a generic or gisty summary of C to be available.

Sperling's subjects were presumably in high-level visual states at T1 that represented individual letters as having specific properties like *having an "A" shape* and *having a "T" shape*. That is, at all events, the best joint explanation of (i) the fact that the subjects could recall the letters in any row after the tone had sounded, and (ii) the fact that at time T2 the subjects recalled that they had *seen* the letters at T1. But only four of those specific properties could be represented in working memory. It wasn't possible for all of them to be represented there. It follows that the original perceptions of the full matrix by Sperling's subjects were not access conscious, according to the understanding of access consciousness that best fits the motivation for having a concept of access consciousness, even though the gists of their perception were fully available to working memory.

There is nothing wrong with *supplementing* Block's original, full-strength concept of access consciousness with a less demanding concept based on gist. Indeed, it might prove theoretically useful to have such a concept available. But the concept based on gist cannot *replace* the full-strength concept. The distinction between full cognitive accessibility and partial accessibility is important in cognitive science and also in the metaphysics of mind. Block's full-strength concept corresponds to significant theoretical needs that the other one cannot fulfill.

In his early formulations of the overflow argument, Block maintained that Sperling's experiment indicates that *phenomenal consciousness* is independent of accessibility. (This is not true of his recent formulation in chapter 1 of Block 2022.) I should emphasize that I have maintained only that the experiment indicates the independence of *perceptual phenomenology*.

Perhaps it would be useful to underscore this point by reviewing the present version of the argument. I have assumed that Sperling's subjects remembered their original perceptions of the full matrix in terms of how the matrix *looked to them*. (This assumption is forced on us by the subjects' claim to have *seen* the full matrix.) Given this background assumption, and given also that Sperling was able to find powerful evidence that the subjects' memories were veridical, it is likely

that the subjects did in fact earlier enjoy phenomenally individuated perceptual representations of the full matrix of letters. The rest of the argument just consists in invoking the definition of access consciousness, and in pointing out that most of the subjects' phenomenally individuated representations did not meet the terms of the definition.

Note that this summary of the argument is concerned exclusively with *phenomenology*. There is no mention of *phenomenal consciousness*. That is, it is neither asserted nor presupposed that phenomenology does or even can possess a third form of consciousness that is independent of introspective consciousness and access consciousness. The argument is simply designed to show that phenomenology is independent of access. The question of whether there is a third form of consciousness that stands in some intimate relation to phenomenology will be taken up later.

In addition to presenting the overflow argument, which is based on an experimental tradition that began in the 1960s, Block has maintained that several more recent experiments in vision science point to the separability of visual phenomenology and access consciousness. In different ways, these experiments show that perceptual activity in the visual areas in the back of the brain can in certain special circumstances occur without being accompanied by cognitive activity that is supported by structures in the front (Pitts et al. 2014, Brascamp et al. 2015). The reader is referred to Block 2022 for exposition of several of the experiments, and for arguments aimed at establishing their relevance to the separability issue.

V. The Categorical Base Hypothesis

We have found that phenomenology does not entail any of the four forms of consciousness that we considered in Section I. At the same time, however, it is evident that (sufficiently robust) phenomenology stands in intimate causal relations to those forms. There is no doubt, for example, that perceptual states with a phenomenal character can provide rational control of action, and have the power to cause episodic perceptual memories and perceptual judgments. Nor is there any doubt that phenomenal states play a role in generating the introspective states that are directed on them. But if relations of this sort are metaphysically contingent, as I have been maintaining, how can they be as intimate as they are?

A natural answer is that robust (that is, highly intense) phenomenology provides a *categorical base* for the causal powers that constitute dispositional introspective consciousness and dispositional access consciousness. (As is usual, I am thinking of the categorical base for a disposition to do X as a set of non-modal properties that are capable, given the laws of nature, of causing X. (For a rich discussion of categorical bases, see Choi and Fara 2018.)) To see the merits of this

view, notice that phenomenology seems to confer a selective advantage on the creatures that possess it. In general, the creatures that possess it are able to provide for their needs more successfully, and in a much wider range of contexts, than creatures like Aplysia, which, given the results of comparative neuroanatomy, pretty clearly lack it. Now, if this is right, then robust phenomenology has biological value. But what does this value consist in? The natural answer is that it consists in the downstream effects of phenomenal states, for, in general, it is what a biological item can *do*, its contribution to the modus operandi of the organism, that the biological value of the item consists in. But it could not be true that the enduring value of phenomenology consists in its downstream effects unless the dispositions to produce those effects were grounded in phenomenology. Since we have seen (in Sections II and III) that the dispositions in question are metaphysically independent of phenomenology, this "grounding" relationship is best described by saying that phenomenology provides the categorical base for the dispositions. (More precisely, it provides part of a categorical base. The total categorical base includes working memory and the various downstream cognitive agencies that working memories of phenomenal states can activate.) (The categorical base hypothesis has also been endorsed by Jesse Prinz. For further discussion, see chapter 3 of Prinz 2012.)

It is part of this view that robust phenomenology also supports the occurrent causal relations that constitute the non-dispositional forms of introspective consciousness and access consciousness.

It probably goes without saying, but I will say anyway, that the categorical base hypothesis does not claim that robust phenomenology is causally sufficient for *all* types of high-level cognitive activity. For example, a perceptual state representing an array of marks on paper may be sufficiently robust to support such cognitive operations as classifying the array as containing letters of the English alphabet and remembering the array as having this property; but it might still be the case that the letters are too small for a subject to read them easily, or too closely packed together for a subject to count them. Evidently access consciousness comes in degrees, where degrees are determined by the number and kinds of cognitive activity that are found in a given case. The categorical base hypothesis claims that robustness of phenomenology is sufficient for a certain base level of access consciousness, but it allows that factors other than robustness may be involved in supporting higher degrees of access.

VI. Phenomenal Consciousness

As I said at the outset, it is widely acknowledged that perceptual states can possess the two forms of introspective consciousness that we have considered and also

the two forms of access consciousness. I will turn now to consider whether there is another form of consciousness that should be acknowledged. This inquiry will occupy us for several sections.

In addition to recognizing introspective consciousness and access consciousness, many philosophers maintain that perceptual experiences also possess a fifth form of consciousness, a form that is independent of occurrent and dispositional relations to second-order representations and to high-level cognitive faculties. This fifth form is often referred to as *phenomenal consciousness*—a designation that is meant to highlight the relationship between it and phenomenal character, which is thought to be more intimate than the relationships between phenomenal character and the other four forms of consciousness. (See, e.g., Block 1995, Chalmers 1997, and Block 2002.)

I will eventually argue that this view is correct, but the path to a full defense of the view is necessarily long and circuitous. For example, some care is needed to distinguish interesting forms of the view from the uninteresting conception of phenomenal consciousness that is suggested by the following passages from a book by James Mill:

> Having a sensation, and having a feeling, are not two things. The thing is one, the names only are two. I am pricked by a pin. The sensation is one; but I may call it sensation, or a feeling, or a pain, as I please. Now, when, having the sensation, I say I feel the sensation, I use only a tautological expression: the sensation is not one thing, the feeling another; the sensation is the feeling. When, instead of the word feeling, I use the word conscious, I do exactly the same thing, I only use a tautological expression. (Mill 1869, 224)
>
> It was of great importance, for the purpose of naming, that we should not only have names to distinguish the different classes of our feelings, but also a name applicable equally to all of those classes. This purpose is answered by the concrete term Conscious; and the abstract of it, Consciousness. (Mill 1869, 225)

With just a bit of anachronism, it is possible to summarize the semantic claim that Mill is expressing here by saying that the adjective "conscious" is a general term that stands for all and only mental states that have phenomenal character. A companion view is that the noun "consciousness" stands for the property *having some phenomenal character or other*. Suppose these claims are true—that there really are uses of the terms in question of the indicated sort. On this assumption, the thesis that there is a form of consciousness other than introspective consciousness and access consciousness is uninteresting. It is tantamount to the thesis that phenomenology is different than introspective consciousness and access consciousness, and this second thesis is truistic. It would be conceded even by those mistaken authors who deny the autonomy thesis.

I hasten to add that there is room for doubt that Mill's semantic claims are correct. Perhaps he would have withdrawn them if he had noticed that there might be minimally robust forms of phenomenology of the sort suggested by the foregoing Leibnizian continuity argument. I have cited Mill's claims because I feel that many people have been tempted by a Millian view. (Certainly I have been.) I will not, however, be placing any dialectical weight on the claims.

The interesting thesis is that there is a use of "phenomenal consciousness" on which it is not pleonastic. Or, to set aside semantics for metaphysics, the interesting claim is that there is a form of consciousness X such that (i) X is distinct from both introspective consciousness and access consciousness, (ii) X is distinct from the property *having some phenomenal character or other*, and (iii) phenomenology + X = phenomenal consciousness. I will use the technical term "*p*-consciousness" for anything that satisfies these conditions.

VII. Proposals about *P*-Consciousness

There are three possibilities as to what *p*-consciousness might be. It might be a property bestowed on phenomenal states by virtue of their relationship to certain other things, or it might be an entity of some kind that serves as a ground or foundation for phenomenal properties. That is, it could be a substance that instantiates them or a property or relation of which they are modes. Further, if *p*-consciousness isn't a relational property of phenomenal states or an entity that serves as a foundation for phenomenal properties, it might be a monadic property of phenomenal states other than the property *possesses some phenomenal character or other*. I will call these possibilities the *relational* view, the *inherence* view, and the *monadic* view. I will discuss a specific proposal from each of these categories, beginning with the first. I hope the reader will feel that the sample proposals I have chosen represent the best hopes for success of their respective categories.

Example of the relational view: Some philosophers maintain that there is a form of consciousness that is similar to introspective consciousness in that it involves second-order representations of mental states, but is different in that it lacks the other features of introspective awareness (Lycan 2004, Rosenthal 2005, Gennaro 2012). According to these philosophers, this form of consciousness does not involve attention, it does not involve working memory, and it is not controlled by mechanisms that determine whether second-order representations are task relevant. In other words, instead of occurring only in special circumstances, as is the case with introspective consciousness, it comes into existence spontaneously with great frequency. More concretely, its constitutive second-order representations are generated automatically whenever the mind is in a sufficiently full-bodied phenomenal state. To give it a name, the view of these philosophers is the

second-order theory of p-consciousness. They urge that the second-order theory is much better equipped than theories of consciousness based on introspection to explain fact that so many of our mental states count as conscious. There is an unbroken stream of conscious states throughout the day, and the second-order theory can nicely explain this fact; but introspection occurs too rarely for theories based on introspection to accommodate it. According to the theory, these conscious states count as conscious because they share the property *being an object of a second-order representation*.

Example of the inherence view: According to the *adverbial theory of p-consciousness*, *p*-consciousness is a generic, undifferentiated relation of awareness, and phenomenology consists of a range of phenomenal modes or determinations of that relation. The view implies, for example, that to be phenomenally conscious of an object O as blue is to be "bluishly" conscious of O, and that to be conscious of O as spherical is to be "spherically" conscious of O. *P*-consciousness is said to be the relation that is common to bluish consciousness of objects, spherical consciousness of objects, and so on. This theory derives from the *adverbial theory of perception* that was originally proposed by Ducasse (1952), and was strongly advocated by Chisholm (1957), among others. (We considered Ducasse's theory of perception in an earlier chapter, but the evaluation of it there was quite brief. The criticisms that I will offer below of the adverbial theory of *p*-consciousness also apply to Ducasse's proposal about perception.)

Example of the monadic view: The *categorical base theory of p-consciousness* begins by reaffirming the categorical base hypothesis that we considered in Section V. That is, it claims that phenomenology that is especially robust serves as a categorical base for the causal powers that constitute introspective consciousness and access consciousness. Here robustness is understood to be like the property of being loud, which is possessed by some noises and not by others. After making this initial claim, the categorical base theory goes on to assert that this sort of highly robust phenomenology counts as conscious in virtue of its role in supporting the various dispositions and causal relations that figure in these other forms of consciousness. It is a form of consciousness because it supports and explains these other two forms.

In summary, we can distinguish three general ideas about the nature of *p*-consciousness, the relational view, the inherence view, and the monadic view. There are reasonably plausible concrete proposals that belong to these categories. All three of these proposals insist on the metaphysical autonomy of phenomenology, but they diverge in their positive claims about the relationship between phenomenology and consciousness. They maintain, respectively, that *p*-consciousness is a relational property that phenomenal states have in virtue of being the objects of second-order representations, that it is a generic, undifferentiated relation of awareness, foundationally involved in relational facts of the form *X is F-ishly conscious of O*, and that it is a type of phenomenology that serves as the categorical

base for some very interesting causal powers. As far as I know, this set of views includes all of the concrete proposals belonging to our three broad categories that enjoy any degree of prima facie plausibility.

I will argue that the first and second of these ideas are dead ends, but that the third is probably valid.

VIII. The Second-Order Theory of *P*-Consciousness

As we observed, the second-order theory has the virtue of being able to explain why so many of our mental states are conscious. It is also true, however, that its explanation is ontologically inflationary: for every conscious state M, the theory posits a representation of M. This would not be a flaw if there were no other way of explaining the broad distribution of consciousness, or if there was independent evidence that the posited second-order representations exist. But it is a serious flaw if neither of these conditions is fulfilled. I will argue that they are not.

First, the broad distribution of consciousness can be explained in terms of access consciousness. As we saw, it is plausible that robust phenomenal states are the categorical bases of the tendencies that constitute dispositional access consciousness. They are also constitutively involved in occurrent access consciousness. It follows from this that robust phenomenal states generally enjoy access consciousness. Hence, the explanatory power provided by the second-order theory is unnecessary—assuming, of course, that second-order awareness is distinct from access consciousness, as its proponents maintain.

Second, as far as I can see, there is no independent evidence that the representations posited by the second-order theory exist.

There is no introspective evidence—in my case at least. When I introspectively examine my conscious perceptual experiences, I am aware of phenomenology, but I have no impression that the phenomenology is accompanied by a second-order representation. Certainly there is no second-order representation with a distinct, proprietary phenomenology. There is no flashing or halo, no pointer or spotlight, that accompanies perceptual phenomenology but is not part of it. Nor is perceptual phenomenology duplicated, in a way that might suggest some sort of mirroring relation. Moreover, most of the time, there are no conscious thoughts that explicitly refer to the phenomenology. When introspection goes looking for second-order representations of perceptual states, it comes back empty handed.

Nor does folk psychology provide any grounds for embracing the second-order theory. Thus, for example, folk psychology clearly allows us to entertain the hypothesis that certain animals have conscious perceptual states but lack the ability to form second-order representations of them. This sort of claim is taken seriously by primatologists (Cheney and Seyfarth 1990, 2008). It can't be ruled out a priori by pointing to some more fundamental folk law. Nor, for that matter, can

the hypothesis that the energy costs would be too great for adult human beings to constantly form second-order states. Second-order representations of phenomenology are as invisible to folk psychology as they are to introspection.

Is there experimental evidence for second-order representations? Lau and Rosenthal argue that there is (Lau and Rosenthal 2011), but, in building their case, they appeal again and again to verbal reports of perceptual experiences, maintaining that they are indicative of second-order representations. This reliance on verbal behavior poses a problem, given that they want to distinguish second-order awareness sharply from access consciousness; for, as we have seen, the ability to issue verbal reports is part of access consciousness. It seems that the methodology of these authors is unable to pry second-order awareness loose from access.

This reflection leads to a dilemma. Either there is behavioral evidence for second-order representations or there isn't. If there is, then the relevant form of behavior is evidence that the representations are access conscious, for accessibility is required for behavior that is complex enough to serve as strong evidence for second-order representations. If there isn't behavioral evidence, then we must look to neuroscience for evidence of second-order representations. But in the absence of introspective and behavioral evidence for them, it is hard to see what grounds we could have for interpreting a particular type of neural activity as second order. Even if we could find a type of activity in the prefrontal cortex that always accompanied a certain type of activity in a perceptual system, why should we construe it as a second-order representation? After all, we lack independent evidence that it even *might* have that status.

It might be suggested that we should recognize second-order representations because they are needed to perform a vital cognitive function of some sort, a function that perceptual states could not perform on their own. But what could that function be? It would have to be quite important to justify the toil and trouble that would have been required of Mother Nature should she have equipped us with it. So it should be very much in evidence if it exists. As far as I know, however, no second-order theorist has ever explained what that function might be. (They could not respond by saying, "Well, of course consciousness is important," because what is at issue here is whether any importance attaches to the specific sort of consciousness that would come from second-order representations of first-order states.)

IX. The Adverbial Theory of *P*-Consciousness

The adverbial theory posits an abstract relation of conscious awareness and claims that phenomenology consists of modes or determinations of this relation. It identifies *p*-consciousness with the abstract relation.

The problem with the adverbial conception is that it just seems false that phenomenology consists of modes of awareness rather than properties that count as objects of awareness. To be sure, we often describe phenomenology by propositions of the form *Object O looks F to agent A*. The only objects of awareness mentioned by such propositions are values of the variable O. Instead of picking out an object of awareness, F has the role of a modifier of the relation *looks to.* The fact that we rely on this construction lends some support to the adverbial conception of phenomenal consciousness. The support is minimal, however, for we also say things that imply awareness of appearances. As we noticed in Chapter 1, we acknowledge, in both discourse and thought, that we are aware of *how mountains look when seen from a distance*, of *how railroad tracks look as they recede towards the horizon*, and of *how tan things look when they are cloaked in shadow*. Moreover, we must suppose that we are aware of perceptual phenomenology if we are to explain how we have knowledge of its various forms. Surely we know what it is for an object to look blue or look small or look trapezoidal? If we could not be aware of phenomenal colors and sizes and shapes, how could we have knowledge of such things?

It might be held that we don't know about phenomenology by participating in facts of the form *O looks F to agent A* but rather by some sort of second-order awareness of such facts. It is, however, very hard to see what that awareness would consist in. It could not be perceptual, because there cannot be perceptual awareness of perceptual awareness. Nor could it be purely cognitive. One can have thoughts about a type of phenomenology (say, the phenomenology of a creature from a different species) without ever having been experientially aware of phenomenology of that type. In general, mere thinking is incapable of making us presentationally aware of perceptual phenomenology. That leaves introspective awareness. But surely we can be in touch with phenomenology when our introspective faculty is not engaged. Perceptual experience always has a phenomenological dimension (when we perceive an object, we are always aware of the appearance it presents), but introspection is comparatively rare. Phenomenology is experienced whenever we open our eyes. Introspection requires an additional faculty and additional energy.

This is not to deny that we can be aware of phenomenology by introspection. It's just that we need a theory of introspective awareness of phenomenology that is compatible with saying that we grasp phenomenology perceptually. There is growing support for such theories in the literature. Suppose a subject S is engaged in becoming introspectively aware of how a certain object O looks. A number of authors (e.g., Harman (1990), Byrne (2018), Gupta (2019), Wu (2022)) maintain that S will proceed in a world-to-mind fashion. Adopting this idea, I will here view S's procedure as involving three steps. (i) S will begin by using attention to boost the processing that determines the appearance of O, thereby enhancing the crispness and detail of the appearance, and also prolonging the experience of the

appearance so that it is easier for S to conceptualize and describe it. (ii) S's second step will be to form such a conceptualization of the properties of O under the guidance of the appearance, using concepts like *blue*, *spherical*, and *smooth*. (iii) Once such a conceptual representation of O is available, S will then form a judgment like *O looks blue, spherical, and smooth to me*. In this process, there is a direct, experiential grasp of phenomenology only at the first stage, which is perceptual. Introspection merely intensifies perceptual awareness of phenomenology and adds a layer of conceptual awareness.

In my view, this picture of introspection is strongly preferable to the principal alternative, which represents introspection as consisting in Russellian acquaintance (Chalmers 1997, Gertler 2011). Russell thought that epistemology requires a primitive cognitive relation between subject and object that is free from inference, interpretation, and all forms of representation, so he took over the commonsense notion of acquaintance and added these further conditions (Russell 1912). But there are a number of mainstream epistemological programs that have no use for this souped-up conception. Moreover, as far as I can see, there isn't a shred of empirical evidence that points to it. Cognitive scientists are baffled when Russellian acquaintance shows up in conversations with philosophers. They think of awareness as being constituted by representations. (As the reader may recall, I had more bad things to say about acquaintance back in Chapter 5.)

In sum, awareness of perceptual phenomenology is a perceptual affair, not a result of introspective awareness of perceptual states. It occurs whenever we are perceptually aware of appearances, which is whenever perceptual experience occurs, not only on those comparatively rare occasions when there is introspective awareness of perception. Moreover, reflection on the nature of introspection shows that the awareness of phenomenology that it provides already existed as perceptual awareness, albeit in a less attentionally enhanced form.

X. The Categorical Base Theory of *P*-Consciousness

The categorical base theory makes two claims. First, highly robust phenomenology serves as the categorical base for the dispositions that are constitutive of introspective consciousness and access consciousness, and also serves as the monadic support for the causal relations by which those dispositions are manifested. And, second, highly robust phenomenal states count as *p*-conscious in virtue of underpinning these dispositions and relations. We have already seen that the categorical base hypothesis is plausible: it is the best explanation of the intimate relation between phenomenology and the other two forms of consciousness. What about the second claim? Is there a reason for embracing it?

Yes. Consciousness plays a role in laws that explain why introspectibility and accessibility are limited in certain ways. Thus, it is a law that only conscious states

can be objects of introspective awareness. It is also a law that only conscious states can fix perceptual beliefs, lay down episodic memories, serve as components of plans, and so on. As we all know, these laws provide valuable explanations of psychological facts of various kinds. For example, they explain why the early phases of perceptual and linguistic processing are not accessible to introspection, and why these processes lack the causal powers that are constitutive of access consciousness. Moreover, they clearly cannot be replaced by similar laws involving introspective consciousness or access consciousness. The principle that only introspectively conscious states can be objects of introspective awareness has no explanatory value. Thus, the foregoing laws of consciousness have a utility and an integrity which indicate that there is a causally important form of consciousness that is independent of the forms we have already recognized.

The laws we have been considering indicate that this form of consciousness can figure in explanations of why various mental states and processes lack introspective consciousness and access consciousness. I now add that it appears that this form of consciousness also figures in laws that can explain the downstream effects of these other two forms of consciousness. Thus, for example, it appears to be a law that conscious perceptual states can, in normal circumstances, fix perceptual beliefs, help to shape plans, and give rise to episodic memories. This law can clearly support highly substantive explanations of these downstream phenomena, unlike the principle that states that access consciousness has these same effects. It is at best a weakly informative explanation of a state S1 to say that it is due to a state S2 that has a disposition to produce states like S1.

It follows from what we have been noticing that this form of consciousness is a locus of causal powers and is therefore eligible to count as a natural kind. But it also follows that it is associated with the same causal powers as a kind that we were considering a few paragraphs back—the kind consisting of all highly robust phenomenal states. This is the best possible evidence that these two kinds are identical.

We have found a strong candidate for the role of *p*-consciousness. I suggest that we should accept both of the following identity statements: "*p*-consciousness = the form of consciousness that figures in the foregoing laws" and "the form of consciousness that figures in the foregoing laws = highly robust phenomenology." Together these identities answer the question we have been considering since Section VI. *P*-consciousness is what occurs when phenomenology becomes highly robust.

To rephrase the points made in this section, it appears that we have a concept of consciousness that we use to explain certain limitations in the scopes of introspective consciousness and access consciousness. We can see this concept at work in laws such as the proposition that only conscious states can be objects of introspective awareness. It is clear that this concept is distinct from, and indeed independent of, the concepts of introspective and access consciousness. If we were to

substitute the latter concepts in the laws, the resulting propositions would be trivialities, altogether lacking in explanatory power. Further, it appears that we use this additional concept in giving deep explanations of the states that figure in the dispositions that constitute access consciousness. Finally, it seems quite appropriate to say that this additional concept is a concept of *phenomenal* consciousness, since it refers to phenomenology that is highly robust.

XI. Conclusion

We can conclude, then, that there are several types of consciousness that perceptual experiences can enjoy: occurrent introspective consciousness, dispositional introspective consciousness, occurrent access consciousness, dispositional access consciousness, and substantial phenomenal consciousness (the property *being phenomenology that is highly robust*). Perhaps there is also a sixth type—Millian phenomenal consciousness (the second-order property *having some phenomenal character or other*). Another conclusion is that phenomenology is metaphysically autonomous, in the sense that it does not entail either of the forms of introspective consciousness and does not entail either of the forms of access consciousness. A more accurate account of the relationship between phenomenology and introspective consciousness and access consciousness is that highly robust forms of phenomenology—that is, *conscious* phenomenology—provide an intrinsic, categorical base for the dispositions and occurrent causal relations that constitute them.

8
Percepts and Concepts

I. Introduction

In this chapter I will be considering one part or dimension of the question of how perception is related to cognition. More specifically, I will be considering perceptual states in relation to high-level cognitive states that are constructed from concepts. This is a very broad domain that includes, paradigmatically, all of the states known as *propositional attitudes*—that is, thoughts, beliefs, desires, intentions, hopes, fears, and so on. How exactly are perceptual states related to these high-level cognitive phenomena? What are the differentiae that distinguish the two realms, and what are the main points of similarity?

I will be arguing that there is a broad chasm separating propositional attitudes and perceptual states. Propositional attitudes play a huge role in cognition, so my arguments will have a broad relevance. I should emphasize at the beginning, however, that there are other parts of cognition in which propositional attitudes are not involved, and that my conclusions will have no bearing on them. Indeed, while I hold, and will attempt to show, that concepts and propositional attitudes are disjoint from perception, I think we must acknowledge that there is an important class of cognitive phenomena that constitutively involve states and processes that are very closely related to perception.

To expand on this point: It is clear that much of cognition consists in the manipulation of visual imagery. When, for example, you are planning a walk in a forest, you may use images of various remembered locations to determine which trail you wish to follow. Equally, you may use visual imagery in deciding what food you would like for dinner, or in deciding how to tilt a piece of furniture in order to get it through a doorway. Further, you may use connected sequences of images—that is, visual simulations—in rehearsing a dance routine, or in predicting how some physical process will unfold (Prinz 2002). Now, as a result of experiments in the 1970s by Shepard and Kosslyn (Shepard 1978, Kosslyn 1980), and a large amount of more recent work in psychology and neuroscience, states involving imagery are now widely believed to be similar to perceptual states. Evidence supporting this view includes parallels between performance on tasks respectively involving imagery and perception (Finke 1989, Kosslyn et al. 2006). It also includes similarities in the neural mechanisms and processing that underlie imagining and perceiving, as determined both by imaging experiments and lesion

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0008

studies (Kosslyn et al. 2006, Broggin et al. 2012). In view of these results, and others to the same effect, it would be a serious mistake to think that all of cognition is independent of perception.

The general picture of cognition I wish to recommend is that instead of being a fully unified, single domain, it consists of several domains that make use of different processes and/or different representational codes (including sentences, maps, images, models, and simulations). The different domains are united by the fact that they are all capable of some form of reasoning, but the forms of reasoning are diverse, so this common feature doesn't amount to a strong similarity relation. Using visual simulation to plan a shopping trip is quite different than going through the steps of a proof in mathematics, and they are both quite different than trying to determine your location on a mental map. (I believe that in taking this essentially pluralist view of cognition, I part company with Ned Block, who, while allowing that cognition can take different forms, seems to hold that manipulation of logically structured propositions holds a position of privilege (Block 2022).)

I add that cognition based on perceptual simulation and manipulation of images is plausibly regarded as evolutionarily more fundamental than forms of cognition that involve concepts and propositional attitudes. The structures at the back of the brain that support perception and imagery are much older than the forebrain structures that support higher-level cognition in ourselves, and it seems likely that our hominid forebears made use of the former structures in short-range planning, social problem solving, inferences about dominance relationships, and other forms of elementary reasoning. After all, our simian contemporaries seem quite adept at such activities (Cheney and Seyfarth 1990, 2008).

This is by way of emphasizing that the ambitions of the chapter are limited: it will focus exclusively on relationships between perception and the dimension of cognition that involves concepts, and will therefore ignore large and historically fundamental parts of the cognitive domain.

To fix ideas, I will begin by sketching a theory of propositional attitudes, taking belief as the principal example, and by sketching an account of concepts. Then I will address the idea, found in philosophers like Quine (1960) and Davidson (1986), that perceptual states have no proprietary representational content, but rather have a kind of vicarious content owing to causal relations linking them to concepts and attitudes. Against this view, I will maintain that perceptual states are metaphysically and even nomologically independent of concepts, and that they can guide behavior and perform other tasks requiring representational content even when they are not accompanied by conceptual liaisons. Then, in the main body of the chapter, I will describe several key differences between perceptual states and propositional attitudes. I will conclude by noting a point of similarity between perceptual and conceptual contents.

II. A Theory of Propositional Attitudes

Belief is a relation between agents and what I will call *Fregean propositions*, where Fregean propositions are complex entities constructed from concepts in accordance with the principles of logical grammar. To illustrate, it is possible to form the Fregean proposition that the Eiffel Tower is brown by combining the concept of the Eiffel Tower, the concept of brown, and the syncategormatic concept *is*, signifying predication. One need only apply a principle of construction to the effect that an atomic Fregean proposition can be obtained by combining a nominal concept with a concept signifying predication and an adjectival concept. All other Fregean propositions are built up from atomic propositions of this sort by applying recursive principles of construction from a finite list. To continue with the illustration, once the atomic Fregean proposition that the Eiffel Tower is brown is in hand, it is possible to obtain the more complex proposition that it's not the case that the Eiffel Tower is brown by applying the principle that if P is a Fregean proposition, so is the result of prefixing the concept of negation to P. Recursive rules of this sort ensure that the set of Fregean propositions is infinite. (This account of propositions is inspired by Frege (1952), but I am not trying to describe or reconstruct Frege's actual views.)

In addition to properties having to do with their constituents and their logical organization, Fregean propositions have truth conditions. I will assume that the truth conditions of propositions are possible arrangements of objects and properties, so that, for example, the proposition that Tibbles (a cat) is on the mat has as its truth condition the arrangement in which Tibbles is located on the mat. There are three reasons for making this assumption. First, the truth or falsity of a proposition depends on how things stand in extra-conceptual reality, and, in particular, on how things stand with the objects and properties that serve as the referents of its constituent concepts. Second, the truth conditions of a proposition include all of the arrangements of objects and properties that *are capable* of conferring truth on the proposition. Even if the actual arrangement happens to confer falsity on a proposition, there might be other arrangements that *would* confer truth on it, if they were actual, and these alternative arrangements are no less pertinent to the truth conditions of the proposition than the actual arrangement. Third, the logical structures of propositions play no role in individuating their truth conditions. Consider two propositions P and Q that are logically equivalent. No matter how much P and Q may differ in logical form, it is guaranteed that if one is true, the other is true as well. It follows that possible arrangements of objects and properties are suitable candidates for truth conditions, for logical structure plays no role in individuating them. Equivalent propositions come out true when evaluated at the same arrangements.

I will also assume that possible arrangements of objects and properties can be modeled by classes of possible worlds.

Given these assumptions, beliefs can be said to have two objects. Where P is any Fregean proposition, the *primary object* of the belief that P is the proposition that P itself, and the *secondary object* is the class of possible worlds in which the proposition that P is true. Both objects are necessary because beliefs have two jobs. One job is to figure in reasoning. This is made possible by the fact that Fregean propositions have logical structures. The other job is to provide guidance as we negotiate a world consisting of objects, properties, and relations. This is made possible by the fact that beliefs have truth conditions that involve such entities.

I have assumed that Fregean propositions have logical structures like those of sentences, and that concepts and propositions, respectively, have referents and truth conditions, just as words and sentences do. In effect, then, I am assuming that concepts and propositions are quasi-linguistic entities. This view is, I believe, widely shared by philosophers, and it also seems to be well entrenched in cognitive science. Anyone who speaks of logical relations among beliefs and attributes truth and falsity to beliefs is committed to some version of the view. I hasten to acknowledge, however, that "quasi-linguistic" is quite vague, and leaves us with a lot of latitude in how we interpret it. The view that Fregean propositions are quasi-linguistic is just the view that propositions are similar to sentences in familiar languages. It does not imply that they are affiliated with any particular spoken language, or any particular classes of languages.

What about concepts? So far I have just said that they are constituents of Fregean propositions, and that their reference contributes to determining truth conditions. Since we have in effect assumed that Fregean propositions have logical structures like those of sentences, we can infer that concepts belong to logical categories. We have already taken notice of four such categories—the category of nominal concepts, the category of concepts signifying predication, the category of adjectival concepts, and the category of logical connectives. As this partial list suggests, logical categories of concepts tend to parallel grammatical categories of words. What about the reference of concepts? Reflection quickly shows that this is a complex topic. Thus, it appears that there are two concepts of reference that can be applied to concepts—an *intensional* concept and an *extensional* concept. Using the intensional concept we can say that it is always true that the concept of β refers to β, regardless of whether β exists. For example, using the intensional concept, we can say that the concept of Pegasus refers to Pegasus. Another distinguishing feature of the intensional concept is that we can only use it to say that the concept of the Evening Star refers to the Evening Star, and that the concept Obama refers to Obama. It doesn't permit us to say such things as that the concept of the Evening Star refers to the planet Venus or that the concept Obama refers to the 44th President of the US. The extensional concept is different in both of these respects. First, we cannot use it to say that the concept of β refers to β

unless β actually exists. And second, when we are using it, if α and β are the same object, it is entirely correct to assert that the concept of α refers to β, and that the concept of β refers to α. We aren't restricted to saying that the concept of α refers to α. Both of these concepts of reference seem to answer to important cognitive needs. (For discussions that explain and justify these claims, see Hill 2002 and Hill 2013.)

Concepts can combine to form Fregean propositions without restriction. Thus, even though the concept of a square and the concept of a circle are contraries, it is possible to combine them to form the proposition that an astronaut found a square circle on the dark side of the Moon. We also have the proposition that an apple on a tree in Cheyenne is both fully red and fully green, and the proposition that there is a village and a male barber living in that village who shaves all and only the males living in the village who do not shave themselves. As with all other contradictory propositions, these propositions have the empty class of possible worlds as their truth condition.

The individuation of concepts is fine-grained. Thus, for example, even though it is metaphysically necessary that water is identical with H_2O, the concept of water is distinct from the concept of H_2O. It is also true that the concept of a bachelor is distinct from the concept of an unmarried adult human male, given that the latter concept has four constituents while the former is atomic, even though the statement "A bachelor is an unmarried adult human male" is an analytic truth. I will not be able to say more about the individuation of concepts here. It is clear that concepts are very finely individuated, but the question of why that is so is a complex issue that requires extended discussion. (I propose a theory of individuation for concepts in Hill 2021.)

Atomic concepts are abstract entities that lack internal structure. Thus, it makes no sense to ask how an atomic concept is spelled, or how computational processes make use of the formal properties of an atomic concept. Questions of this sort make sense only when they are concerned, not with concepts themselves, but with the mental representations that *realize* concepts in the internal psychologies of individuals.

I should say more about realization. Concepts and Fregean propositions are abstract entities. In effect, they are types or properties that are realized, and arguably multiply realized, by concrete representations in the minds of individual agents—concrete representations that are generally known as representational *vehicles*. Unlike concepts and propositions, vehicles have proprietary casual powers. But vehicles are similar to concepts and propositions in a number of important respects. Thus, in order to count fully as realizers of concepts, vehicles corresponding to concepts must belong to classes that correspond to logical categories, and they must bear relations to objects and properties in the world that correspond to reference relations. Further, to count fully as realizers of Fregean propositions, vehicles must have logical structures, and they must possess

relations to states of affairs that correspond to truth conditions. In sum, vehicles must honor the concepts and propositions that they realize by having a linguistic character. Otherwise there would be no grounds for viewing the vehicles as realizers. It follows that the vehicles corresponding to concepts in the mind of a subject S must be words in a language that is in some sense possessed by S, and that the vehicles corresponding to propositions must be sentences in that language, or at least sentential fragments. At present, not much is known about the nature of realizing linguistic vehicles. Perhaps they are words or sentences in a widely shared language of thought, but it seems no less plausible that they are affiliated with spoken languages, with different spoken languages providing different vehicles for different subjects. We will have to leave questions in this area open, pending future discoveries by cognitive science.

In addition to realizers of concepts and propositions, there must also be realizers of propositional attitudes, such as the belief that it is raining and the desire to carry an umbrella. According to a familiar picture, the realizers of attitudes involve a set of functionally defined "boxes"—the belief box, the desire box, the intention box, and so on. Sentential vehicles can be inscribed in these boxes, and the full realizer of an attitude consists in the presence of a sentential vehicle in a box. To illustrate, the realizer of the belief that P consists in the presence in the belief box of a sentential vehicle that realizes the proposition that P. The functional properties that define the boxes are computational procedures that take sentences (or perhaps sentence fragments) as inputs and deliver sentences (or sentence fragments) as outputs. The full set of causal powers of the realizer of the belief that P are determined in part by the computational procedures that are associated with the belief box and in part by the words and grammatical structure that individuate the sentential vehicle that realizes the proposition that P.

As noted, this picture is familiar. It also enjoys a wide popularity. I will assume here that it is correct.

To add a few brush strokes: The computational processes that define the boxes are assumed to correspond to processes involving propositional attitudes that are specified by folk psychology and cognitive science. To illustrate, folk psychology implies that beliefs and suppositions figure in deductive reasoning, and, correlatively, it is part of the boxology picture that the belief box and the supposition box are partially defined by computational processes that correspond to the various types of deductive inference. Further, to say that a proposition-vehicle is "in" one of the boxes is just to say that it is available for use by the computational processes that define that box.

There is one more issue involving propositional attitudes and their realizers that should be discussed. As we have been observing, the propositions that serve as objects of human attitudes share some important features with sentences, and, as a result, it is natural to suppose that the concrete vehicles that realize them in the minds of individual human agents are linguistic entities. But this leaves us

with the question of how to view the attitudes of non-linguistic creatures like apes and monkeys. Ethologists have found reason to attribute beliefs to such creatures, but there appear to be few grounds for crediting them with inferences involving beliefs with complex logical structures (Cheyney and Seyfarth 1990 and 2008). This fact can be interpreted in various ways, and adjudicating among these interpretations is a complex task. Here I will just point out that the fact doesn't preclude the hypothesis that animal beliefs have Fregean propositions as their objects. On the contrary, the fact can be accommodated by supposing that certain animals have beliefs with Fregean propositions as objects, but are only capable of believing a very limited range of such propositions—perhaps only monadic propositions of the form *a is F* and relational propositions of the form *a R b*. As far as I know, this idea is consistent with the available data.

If we adopt the idea, what view should we take of the realizers of the simple Fregean propositions that serve as the objects of animal beliefs? Clearly, we cannot suppose that the realizers are sentences in a language spoken by the relevant believers; but it is not hard to imagine that these believers have fairly abstract representations of individual objects, fairly abstract representations of a limited set of properties and relations, and the capacity to combine the former representations with the latter in structures that amount to predication. Here "abstract" might mean no more than that they can be deployed in thought independently of stimulus control. (See Cheney and Seyfarth 2008 and Camp 2009 for discussion of how the beliefs of a baboon concerning relations of dominance among members of its troop are realized by quasi-linguistic vehicles.)

To be sure, it can be tempting to attribute logically complex beliefs to animals; and if we were to yield to that temptation, it would be difficult to maintain that beliefs generally have linguistic or quasi-linguistic realizers. To experience the temptation, recall that in times of drought, a grandmother elephant may lead her herd several dozen miles to find water. Seeking an explanation for this behavior, one might credit the elephant with a belief to the following effect: if we follow such and such route, we will arrive at place where water is abundant. A linguistic or quasi-linguistic realizer for such a belief would have to have a conditional form, and perhaps also an existential quantifier attached to a representation of water. When we reflect, however, we appreciate that there are alternative ways of explaining the behavior that are more parsimonious, since they credit the elephant with fewer high-level cognitive abilities. For example, we might suppose that the elephant is guided by episodic memories linking water to distinctive environmental features, and by a map consisting of memories of landmarks and dead reckoning computations from previous journeys. To give another example, we might find it convenient to assert that a chimp Alf believes that if he grabs a certain banana, the alpha male Fred will attack him, but this might signify no more than that Alf has a disposition to move from visualized simulations of food theft to images of Fred on the warpath.

I conclude this discussion of propositional attitudes and concepts by recommending a realistic attitude concerning their existence and objectivity. Quine and other philosophers have argued for anti-realism, while still others have maintained that the reality of attitudes and concepts is in some sense relative to the interests or explanatory practices of human beings. As I see it, these positions are called into question, if not decisively refuted, by the fact that cognitive science has taken over the explanatory apparatus consisting of propositional attitudes and concepts from folk psychology and put it to use with impressive success. To be sure, there are many unanswered questions about the apparatus, and attempts to throw light on certain parts of it (such as the basis of the individuation of concepts) have turned out to be frustratingly difficult. But whether we are justified in believing a theory, and therefore justified in taking a realistic attitude toward the entities in its domain, is largely determined by the theory's predictive and explanatory success, and it is clear that the theory of attitudes and concepts has achieved high marks in both areas.

III. Percepts are Metaphysically Independent of Concepts

A number of philosophers have believed that concepts figure constitutively in perception, and that the contents of perceptual states are determined by their conceptual constituents. The list includes Kant (2003) and Wittgenstein (1953), in earlier times, and more recently McDowell (1996), Fodor (2007), Quilty-Dunn (2016), and Mandelbaum (2017). When we consider the question, however, we find that there are several quite substantial arguments for the opposing view. I will review two of them.

First, there is the argument that perceptual representations of sensible characteristics can achieve a fineness of grain that is denied to conceptual representations of the same domains (Evans 1982). The argument is usually developed in terms of colors. There are disagreements concerning the number of colors that adult humans can discriminate, but it will serve my purposes here simply to choose one of the lowest estimates. I will assume that the number is 1,000,000. How many color concepts are there? One way to estimate this is to ask how many words there are in the vocabularies of people who speak the English language. The answer is that the sizes of most vocabularies fall in the 20,000–35,000 range. Since the portion of a person's vocabulary consisting of color words is inevitably much smaller than the vocabulary as a whole, it appears that the number of visual representations of colors is many times greater than the number of names for colors. It is reasonable to conclude that this also holds for nominal concepts of colors.

This takes us partway towards an interesting conclusion, but we must also consider whether it is possible to match the expressive power of visual representations of colors by forming descriptions of them. The answer to this is that it is in

principle possible, in some very thin sense of "in principle," to specify every color by a description, but that in actual fact we lack the resources to achieve this. More specifically, we are unable to add all of the descriptions to our working conceptual vocabularies. Thus, for example, it is in principle possible to capture every shade of orange by a description of the form "shade of orange that is X discriminable shades distant from pure yellow," or by a description of the form "shade of orange that is Y discriminable shades distant from pure red," but there is no one on Earth who actually deploys such descriptions, and this seems to be the result of natural law. To actually apply a relevant description to a given shade of orange, it would be necessary to fix upon the exact distance of the shade from unique yellow (or its exact distance from unique red). This would require an extensive measuring procedure. No one could do it on the fly, if indeed it could be done at all. Accordingly, it would be ludicrous to claim that such descriptions are partially constitutive of our everyday color experience.

Names and descriptions provide two ways of picking out colors conceptually, but it is also possible to specify them by demonstratives of the form "*that* shade of color." Such demonstratives must be equivalent to perceptual representations in expressive power. After all, for every shade of color *C*, it is in principle possible to attend to an object that possesses *C* while entertaining a demonstrative concept of the given form. This is all quite true. It is clear, however, that it has no tendency to show that demonstrative concepts figure constitutively in perception. On the contrary, it is evident that one must perceptually represent a shade of color in order to be able to demonstrate it. Perceptual demonstratives depend constitutively on independently existing perceptual representations for their contents. This can be seen quite easily by considering the corresponding linguistic case. If you were to utter the phrase "that shade of color" in a context in which you were surrounded by different colors, and while you were attending perceptually to something other than a color—say, the song of an unseen bird—the phrase would be empty. Conceptual and linguistic demonstratives are parasites, not autonomous cognitive equipment. (I find it surprising that this rather obvious point receives so little acknowledgment in the literature.)

It is sometimes suggested that we have recognitional abilities involving all the shades of color, and that the ability to recognize a shade of color is in some sense conceptual in character (McDowell 1996). As far as I can see, however, both of these claims are misguided. According to the *OED*, recognition consists in identifying a currently perceived item as one that has been perceived before. Accordingly, an ability to recognize instances of a property presupposes a previous encounter with an instance of the property, and also a memory of that instance; but it is clear that no one has encountered all of the shades of color that human beings can discriminate, and, anyway, it seems very unlikely that we retain memories of all of the shades of color that we encounter. Moreover, when a memory of a previous encounter with a given shade of color *does* exist, it presumably is

a perceptual memory, existing independently of conceptualization, and recognition presumably involves matching the currently presented shade of color to that perceptual memory. What then could be the basis for saying that recognitional abilities are conceptual? To be sure, the possession conditions for some concepts may involve recognitional abilities, but that shows only that recognitional abilities may in some cases be constitutive of concepts, not that concepts are implicated in recognitional abilities.

Another point, emphasized in Block 2022, is that memories of perceptual states are less fine-grained than perceptual states themselves. Even if one has witnessed a specific shade of, say, blue, it is unlikely that one has the ability to recognize that specific shade in future encounters with it.

It might be urged in response that the claims about recognitional abilities should be reformulated as claims about discriminative abilities, and that when so revised, they become more plausible. But this seems quite wrong. Discriminative abilities are pretty clearly non-conceptual. It is necessary to discriminate between shades of color before we can think about them.

A second argument that perceptual content is not determined by conceptual content has to do with the complexity of perceptual experience (Hill 2009). We can be perceptually aware of complex arrangements that we are unable to name. This is true, for example, of the distributions of color on carpets, and also of textures and shapes. There is no name for the particular pattern of curved and slanted lines we see when we look at someone's hair. But is it always possible to capture complex arrangements in descriptions? No. Suppose you are looking at a distant mountain range and you attend to the jagged line formed by the peaks and ridges, all of which have proprietary heights. Using your system of visual representation, you are able to take in this line at a glance. But what would you have to do to describe it? It would be necessary to draw or photograph the line, impose an extremely fine-grained co-ordinate system on the resulting image, and then extract an equation for the line using analytic geometry. It is clear that we do not deploy, or even possess, descriptions of this sort for the shapes we encounter in everyday life. In the case of the distribution of colors on a carpet, you could proceed without first drawing or photographing it. You could instead drop an extremely fine-grained plastic grid on the carpet, and then specify, for each cell of the grid, what shade of color is present there. But this case is like the preceding one in that it requires special equipment, and it is also like the preceding one in that it would require a great deal of time to deploy the equipment and construct the description. Perceptual experience differs sharply from conceptualization in its ability to compress complexity into a single momentary representation.

As with the fineness of grain argument, this complexity argument has to address the point that there is a sense in which conceptual demonstratives have the same expressive power as experiential representations. But the answer is the same. The ability of perceptual demonstratives to capture complexity is constitutively parasitic

on the corresponding perceptual ability. Further, any attempt to answer the complexity argument by appealing to recognitional or discriminative abilities would be vulnerable to objections like the ones we considered in connection with the fineness of grain argument.

As far as I can see, the two lines of thought we have been reviewing count decisively against the idea that perception constitutively involves conceptualization, but there are other compelling arguments in the literature, several of which may be found in Block 2022. Appealing to robust bodies of experimental work (Franklin et al. 2014, Skelton et al. 2017, Miceli et al. 2001, M. van Zandvoort, T. Nijboer, and E. de Haan 2007), Block plausibly maintains that infants have a full range of color experiences while altogether lacking color concepts, and that the same is true of victims of a conceptual disorder known as *color agnosia*.

I have been concerned in this section only to deny that concepts are constitutive of perception. This is of course perfectly compatible with acknowledging that conceptually structured states can exert a causal influence on perceptual processing, for example by directing perceptual attention toward objects and properties that are relevant to current tasks. We know that such causal influence exists, though it seems unlikely that it is pervasive (Firestone and Scholl 2016, Block 2022).

IV. More Differences Separating Percepts from Concepts and Propositional Attitudes

In this section I will review five additional differences between belief and perception. Together they will help us to appreciate how perception contributes distinctively to human survival and flourishing.

1. The most obvious of the additional differences is that perception has the job of acquiring information about the world and subjecting it to initial processing, while concepts and propositional attitudes have the job of encoding that information in a format that makes it available for problem solving, explaining, predicting, planning, decision-making, evaluating, and other high-level cognitive activities. Correlatively, perception is proximally dependent on external causes, while the dependence of concepts and attitudes on external causes is distal, being thickly mediated by perceptual processing and cognitive processes.

2. A second signal difference involves logical properties and relations. As we observed earlier, it holds quite generally that concepts belong to logical categories and that Fregean propositions have logical forms. Indeed, since Fregean propositions are generated by the recursive rules of logical grammar, they can have logical forms that are arbitrarily complex. Perceptual states have a few features that approximate logical properties, but the features in question are different than the logical properties possessed by beliefs, and their involvement in perception is much more modest in its proportions.

Before we begin to discuss this issue we should take note of an important distinction. Let us agree to say that a perceptual *judgment* is an occurrent belief that is proximally caused by a perceptual state, and that receives very strong epistemic support from the state. Perceptual judgments take Fregean propositions as objects, and it is therefore trivially true that they are individuated by logical structures. They are typically expressed by terms like "sees that." The question before us here is not the uninteresting question of whether perceptual judgments have logical structure, but the much more difficult one of whether the perceptual states that cause and justify perceptual judgments have logical or quasi-logical structure in their own right.

Let's begin by considering the simplest case, that of conjunction. It's clear that we can perceptually represent an object as being both round and red, and that we can also perceptually represent a shelf as containing both a tall vase and a short one. It follows that perception can register certain conjunctive relationships. It seems, however, that the form of conjunction that is involved in such representations is not the same as the form that can figure in Fregean propositions, for, as we saw, the latter is quite promiscuous. To illustrate, conceptual conjunction can be combined with the concept of a square and the concept of a circle to produce the contradictory Fregean proposition that there is an object that is both a square and a circle. There is no comparable configuration among perceptual states. Equally, although it's possible to form the contradictory Fregean proposition that there is a shelf containing two vases such that the first is taller than the second and the second is taller than the first, there is no comparable perceptual representation. In general, the representational systems used by perception lack the wherewithal to represent contradictions. This places a substantial constraint on the conjunctions it can express. (Alternative versions of this line of thought are put forward in Block 2022 and Pautz 2021. Here is Pautz's version of the main claim: "An individual cannot experientially represent a single surface as having two distinct pure colors, such as pure red and pure blue. Likewise, an individual cannot experientially represent the same object as having distinct shapes, such as round and square" (Pautz 2021, 131).)

It is sometimes maintained that figures like the Penrose triangle (Figure 8.1) and the Escher waterfall show that the visual system can represent impossibilities, but this seems wrong to me (Penrose and Penrose 1958). As I see it, as far as purely visual experience goes, we never represent an impossibility, only the possible parts of an impossibility.

To be sure, it is possible to take all of the parts of the Penrose triangle in at once, but then we are only glancing at it in a superficial and unfocused way, perceiving it as a two-dimensional image of a triangle, without appreciating the three-dimensional implications of the suggestions of occlusion and the direction of the light source that artists cleverly impart to the image. One must attend successively to the corners of the triangle, appreciating the three-dimensional import

Figure 8.1. The Penrose Triangle.

of each one, and then attempt unsuccessfully to combine them into a unified percept, before realizing that the figure is impossible. That realization is a purely conceptual aperçu.

It appears, then, that perceptual representations can signify a type of conjunction that is weaker than conceptual conjunction; but how do they accomplish this? What formal or syntactic properties of perceptual representations endow them with this expressive power? A natural suggestion is that the formal counterpart of conjunction is predicate binding—the operation that binds two or more representations of properties together to form a representation of a single object (Block 2022). This idea provides an attractive explanation of how it is possible to form perceptual representations that are counterparts of Fregean propositions like *there is an object before me that is both red and round*. Moreover, we have some understanding of the neural processes that underlie predicate binding: they probably consist in synchronization of the firing patterns that represent the conjoined properties (von der Malsburg 1999, Engel et al. 2001, Goldfarb and Treisman 2013).

So far so good. But this explanation of perceptual conjunction works only in the case of predicate conjunction. It can't be extended to explain how representations of states of affairs involving distinct objects can be conjoined. Consider two representations that respectively express the state of affairs *there is a thin vase before me* and the state of affairs *there is red book before me*. Clearly there is a representation that expresses the conjunction of these two states of affairs—viz, the state of affairs *there is a thin vase before me and there is a red book before me*. But we can't account for this third representation in terms of the hypothesis that perceptual conjunction consists in synchronous firing. For suppose otherwise. Then there is a perceptual representation R such that (i) R is formed from the representations of the state of affairs *there is a thin vase before me* and the state of affairs *there is a red book before me*, and (ii) all of the components of R fire synchronously with each other. In particular, the perceptual representations that

respectively express thinness and redness fire synchronously, and both fire synchronously with the perceptual representations that respectively express the property of being a vase and the property of being a book. It follows that both thinness and redness are predicated of the vase and both are predicated of the book. But this means that, contrary to hypothesis, R would not express the state of affairs *there is a thin vase before me and there is a red book before me*, for in this state of affairs, only thinness is a property of the vase and only redness is a property of the book. In sum, if conjunction is a matter of synchronous firing in the case of representations of states of affairs as well as in the case of representations of properties, it is impossible to express the given state of affairs by conjoining the representations of *there is a thin vase before me* and *there is a red book before me*.

Instead of saying that conjunction is a matter of synchronous firing in the case of representations of states of affairs, we should just say, I suggest, that representations of states of affairs are linked by the fact that they occur at the same time, as parts of a single larger perceptual representation. (Note that occurring at the same time is different than having components that fire synchronously.) That is, according to this suggestion, we should allow that I can represent that there is a thin vase in front of me, and also that I can also simultaneously represent that there is a red book in front of me, but we should deny that there is perceptual representation that counts as the conjunction of exactly these two representations. Instead, there is just the fact that they are both *parts* of the all-inclusive visual representation that I token at a particular time. If this is right, then conjunctions of perceptual representations of states of affairs are quite different than conjunctions of perceptual representations of properties. That is, there are two types of perceptual conjunction. Moreover, conjunctions of perceptual representations of states of affairs are quite different than conceptual representations of states of affairs. Conceptual conjunctions cannot be analyzed in terms of temporal co-occurrence or common parthood. (For another difference between conceptual conjunction and perceptual conjunction involving different objects, see Goldfarb and Treisman 2013.)

What about disjunction? Here is what Pautz says about this topic:

> [A]n individual cannot experientially represent merely *that there is either a red square in front of you or a green sphere on your right*, without experientially representing anything more specific. What would that be like?
>
> (Pautz 2021, 131)

Block concurs:

> One can see something as intermediate or indeterminate between red and blue (e.g. purple) but not as having the disjunctive property of simply being red or being blue. Note that something that is indeterminate between red and blue

> does not satisfy the accuracy condition for being red or being blue; to satisfy that disjunctive accuracy condition, something must be red simpliciter or blue simpliciter. (Block 2022, chapter 4)

The point that Pautz and Block are making here can be expressed as follows: Although there is a specific phenomenology associated with seeing something that is red, or at least a specific type of phenomenology (there are many shades of red) and a specific phenomenology associated with seeing something that is green, or at least a specific type of phenomenology, there is no specific phenomenology, or even a specific type of phenomenology, associated with seeing something as either red or green.

This seems quite true; but, to switch cases, couldn't one glance at a figure and register only that its shape is somewhere between a heptagon and a decagon—that is, either an octagon or a nonagon? And, if so, what does that tell us about perception's ability to represent disjunctions? Equally, couldn't one look at the coins on a table and register only that the number of coins is somewhere between ten and twenty? There is no doubt that cases answering roughly to these descriptions occur. Indeed, such cases occur all the time in everyday life, and they are studied in a large literature that sees them as effects of a faculty known as the *analog magnitude system*. But, as I read the literature, the properties that the analog magnitude system attributes aren't precise disjunctive properties like *has between seven and ten sides* or *has between ten and twenty members*, but rather vague properties like *has roughly nine sides* or *has about fifteen members*. If this is right, then cases of the sort we've been considering don't indicate the presence of a disjunctive element in the perceptual representational code. (See, e.g., Beck, 2018.)

Turning now to existential quantification, if the position I described under the name "existentialism" in an earlier chapter is correct, all perceptual states have existential contents, and also have formal devices that function like existentially quantified variables. This hypothesis seems to be the only way of accounting for the functional and normative properties we attribute to hallucinations. We should note, however, that the existential character of perceptual representations is quite different than the existential character of Fregean propositions and beliefs. Thus, there is a Fregean proposition to the effect that there exists at least one rectangle, and it is possible to believe that proposition just as it is, without supplementing it with concepts that attribute additional, more determinate properties to the rectangle. Quantification functions quite differently in the case of perception. Apart from very special cases, such as perception of a Ganzfeld, it is impossible for there to be a perceptual state that represents the existence of an object and attributes exactly one property to it. For example, there could not be a perceptual state that has the content *there being a rectangle* but no other content. A perceptual state would at least have to attribute a color to the rectangle (though perhaps only a

non-chromatic color), and a position in space. It would also have to specify a size and an aspect ratio (though these specifications would be vague if the rectangle was positioned in the periphery of one's visual field). Generally speaking, in the case of perception, existential content must satisfy a certain *repleteness condition*: it must involve properties from a range of different categories. This is not true of beliefs. (Cf. Pautz 2021, 185.)

Once again, then, we have found that a logical concept has a counterpart in the perceptual realm, in the sense that perceptual states have structural features that enable them to express a subset of the contents that the logical concept can be used to express, but that the perceptual counterpart is a pale shadow of the concept. I will now point out an even more striking difference between the logical dimension of cognition and the logical dimension of perception. The latter has no counterpart of negation. (Cf. Block 2022, 117–20.)

What would it be for a perceptual state to express negation—that is, to deny that a state of affairs is actual, or that a property is instantiated? When we reflect, we find, I suggest, that perception can only deny that an object instantiates a property P by representing the object as instantiating a different property that is a contrary of P. Consider, for example, what is happening at the purely perceptual level when you are prompted by a visual experience E to form the perceptual judgment that the substance in your test tube has not turned green. E must have a specific phenomenology of some kind, and it must be this phenomenology that endows the experience with the power to prompt the judgment. But there is no specific phenomenology associated with not being green: an object that does not look green could look blue or look yellow or look greenish blue, or look any one of a host of other colors, but there is no such thing as looking not-green. Accordingly, when E prompts you to judge that the substance in your test tube is not green, it must be that E has a specific positive phenomenology that precludes green, rather than a not-greenish phenomenology. Perhaps the substance looks blue. Or perhaps it looks colorless, by which I mean that the substance might have a positive visual appearance like that of water.

There are, however, philosophers who believe that there is a phenomenology of absence (Farennikova 2013, Sartre 1956). Sartre maintained (1956, 9–11), for example, that when one enters a café expecting to see one's friend Pierre, and finds that he isn't present, one perceives his absence. But could this be true? It seems more natural to suppose that one's visual experience on entering the café is replete with positive phenomenology: one sees chairs, tables, waiters, patrons, and so on, and all of these aspects of one's experience are positive. Where is there room in all of this for the absence of Pierre? Each of the things one sees precludes the presence of Pierre, but seeing things that preclude his presence does not add up to seeing his absence. Evidently, one does not *perceive* his absence. Rather, one must *infer* his absence from the fact that all positive items in the café preclude his presence. But

what is it to be aware that a particular positive item (as opposed to a series of them) precludes a presence? Clearly that too must be an inference from what is perceived. But what form does the conclusion of this inference take? Could it be that one forms a visual image of Pierre that is quite faded, or has a cross through it? Of course not. The conclusion that Pierre's presence is precluded is a judgment, and involves the *concept* of negation. We can think of it as compressing the following little argument: that is a table; Pierre is not table; so that is not Pierre.

We can conclude, then, that perceptual representations have only a fraction of the logical power of conceptual representations. It would be a serious mistake to suppose that these two forms of representation have a common logic.

3. A third difference between conceptual and perceptual representations is that the latter have an *analog* structure. This has been noticed recently by several philosophers, principally Neander (2017), Shea (2018), and Block (2022). Roughly, a representational system is analog if the individual representations that figure in the system are organized in a quality space that parallels a quality space involving the items that the representations signify. To explain this idea more fully, I need to appeal to a domain R1, R2,…, Rn of representations, a domain E1, E2,…, Em of represented entities, and a function *f* mapping the former onto the latter. I also need to suppose that there is a similarity relation $\mathbf{R}_R$ linking R1, R2,…, Rn and a similarity relation $\mathbf{R}_E$ linking E1, E2,…, Em. Then I can say that the system involving R1, R2,…, Rn is analog if and only if, for any two representations Ri and Rk, Ri and Rk bear the relation $\mathbf{R}_R$ to each other just in case *f*(Ri) and *f*(Rk) bear the relation $\mathbf{R}_E$ to each other. Finally, I can extend this basic idea by supposing that R1, R2,…, Rn are arranged in a similarity space (i.e., a quality space) of a number of different dimensions, and by supposing that E1, E2,…, Em are arranged in an multi-dimensional similarity space of their own.

It may be useful to have an example of a quality space before us. Figure 8.2 shows one, due to Paul Churchland (1995, 28), in which faces are the values of E1, E2,…, and En.

This representation of face space is of course artificially simple. It captures only three dimensions along which faces can be said to resemble each other—namely, nose width, eye separation, and mouth fullness. We would need to add a number of other dimensions in order to be able to capture all of the distinctions among faces that our perceptual system is able to register. (It is now known that a fifty-dimensional space would be adequate. See Chang and Tsao 2017 and Tsao 2019.) But Churchland's space nicely illustrates the relevant point, which is that families of external stimuli can be seen as points in n-dimensional similarity spaces. Also, once an origin and co-ordinates have been specified for such a space, it is possible to speak of degrees of similarity, and it is also possible to represent each point X in the space by an n-tuple specifying the distance from the origin to X along each of the n-dimensions.

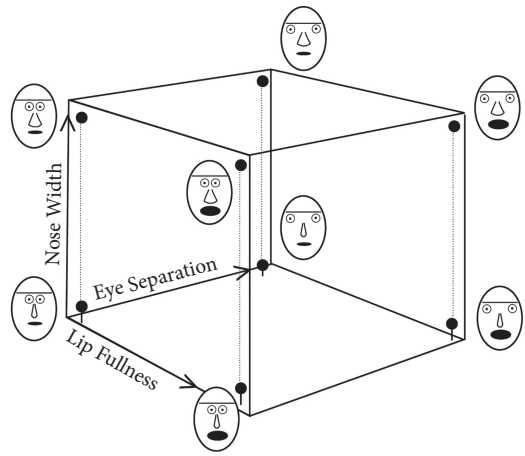

Figure 8.2. A three-dimensional face space.
(Reprinted with the permission of MIT Press from Churchland 1995.)

This gives us a way of imagining what similarity spaces of represented entities are like. What about the similarity spaces of perceptual representations? We can visualize them as n-tuples or n-dimensional vectors in a space determined by n neurons, with each position in a vector corresponding to a rate of firing of one of the neurons, and each neuron corresponding to a dimension in the space of represented entities. (Cf. Shea 2018, 132.) Like Churchland's three-dimensional face space, this visualization is a simplification. Contrary to the visualization, it may require several or even many neurons to create a counterpart of a dimension in the space of represented entities. (It appears to require 205 neurons to capture the fifty dimensions of face space. See Chang and Tsao 2017 and Tsao 2019.) But we need not dwell on this complication here. What matters is that we can think of representations of external entities as n-dimensional vectors consisting of firing rates of neurons, and we can think of such representations as organized into similarity spaces in which distances between vectors is determined by similarities among the firing rates that serve as the vectors' components.

This gives us an idea of the nature of analog representation. We must now consider arguments that it actually figures in perceptual representation. Also, we should consider how extensive its role might be.

There is beginning to be a scientific literature that addresses these issues. I have already mentioned a study that provides a compelling proof of concept—Chang and Tsao's 2017 paper showing that all human faces can be located in a fifty-dimensional similarity space and that vectors over 205 neurons can represent all of the points in this space, with similarities among vectors corresponding to facial similarities. But there are other studies with the same moral. To illustrate, Kiani

et al. (2007) showed natural and artificial images to macaque monkeys while using cell recordings to study the resulting patterns of activity in the inferotemporal cortex. They found that similarity relations among neural patterns corresponded to the category structure of the images. Thus, for example, the patterns correlated with images of fish were more similar to each other than to any of the patterns correlated with images of butterflies. (For additional examples and illuminating discussions of their significance, see Shea 2108, 132–7. The reader is also referred to Block 2022 for arguments for analog representation that are different in kind than the one offered here, being grounded in perceptual imagery and perceptual memory rather than perception itself.)

Further, as Karen Neander has emphasized, there is also a general, theoretical reason for supposing that perceptual representation has an analog character (Neander 2017, 200–3). Neander begins by observing that systems of perceptual representation always contain representations of many properties to which agents have not been exposed in acquiring the systems. Consider colors, for example. We are forever encountering new colors, and the same was true of those of our ancestors who existed when selective pressures were shaping human perceptual systems. That is, the representational power of the color system that was acquired by our ancestors went far beyond the exposure to particular colors that our ancestors had. The system they acquired could represent many colors that they had never encountered.

How is this possible? We don't want to say that evolution selects certain representations because they *could* be triggered by certain stimuli, and therefore *could* carry information about them. Evolution doesn't select individual states that haven't proved their worth. But evolution could select a *system* Σ of similar representations that has proved its worth by containing a number of representations that have a history of encoding information about members of a domain D. Suppose that the members of Σ form a quality space that is isomorphic to a quality space involving the members of D, and that this isomorphism is such that the members of D have the causal power to elicit corresponding members of Σ. It makes sense to say that Σ was selected because of the track record of this isomorphism, and it also seems appropriate to say that the so-far unused members of Σ represent the members of D to which they correspond under the isomorphism. This is an extension of the conception of perceptual representation that we arrived at in Chapter 1, but, as we saw there, it is helpful to adopt a pluralistic attitude toward representational capacities.

Here is how Neander makes the point:

> Along the same lines, the proposal that includes [the analog-representation hypothesis] allows that not every determinate in a sensory-perceptual system's representing or represented ranges must have featured—in person, as it were—in

> the function-conferring and hence content-determining conditions for the system. There can . . . be sensible or perceptible properties represented by a system's representations—even by its representational simples—that never featured in the system's function-conferring selection history. Some represented determinates need never have been (nor need ever be) instanced. What is required, in the case of a naturally occurring analog representational system, is that *enough* determinates in its represented and representing domains featured in its history—enough to fix the relevant constraints (i.e., enough to "specify" or select for the relevant analog relations). (Neander 2017, 200–1)

In speaking of "function-conferring and hence content-determining conditions for the system," Neander is invoking her view that a perceptual representation R has an environmental property P as its content if R has the biological function of encoding information about P. Her point in the passage is that we can explain how the representations in a system signify properties they've never been exposed to only if we suppose (i) that the contents of most of the representations in the system are determined by similarity relations to other representations that have been exposed to the properties they represent, and (ii) that the items signified by these representations are connected by similarity relations that correspond to the forementioned similarity relations among the representations.

There are, then, some very powerful reasons for thinking that systems of perceptual representation have an analog nature. Assuming that this thought is correct, we have found another respect in which systems of perceptual representation diverge from Fregean propositions and their realizers; for there is no reason to think that either concepts themselves or the internal vehicles that realize concepts stand in similarity relations that correspond to similarity relations among their referents. Indeed, insofar as concepts and their realizers are linguistic or quasi-linguistic in character, there is strong reason to deny that similarity plays a significant role in conceptual representation. Shades of red are more similar to shades of orange than to shades of green, but there are no corresponding facts involving the words "red," "orange," and "green." Generally speaking, similarity plays only a very small role in determining the referents of words. (It cannot be said that it plays no role. For example, if we think of "100," "101," "102," etc. as words, then it is true that words designating numbers in the hundreds are linked by similarity.)

4. I turn now to a fourth difference between conceptual and perceptual representations. The property of being analog is a property of systems of representation. As we have seen, a system of representation has this property in virtue of a correspondence (an isomorphism or a homomorphism) between a quality space involving the individual representations of the system and a quality space

involving the extra-mental items that the representations signify. Analog representation is both similar to and different from *iconic* representation. Instead of being a property of systems of representation, iconicity is a property of individual representations. (At least, this is how I will use "iconicity" here. There is a confusing practice in the literature of using "iconic" and "iconicity" to refer to systems of analog representation.) Now, like analog representation, iconic representation depends on isomorphisms (or homomorphisms) from representations to represented items. The difference is that now the isomorphisms link parts of individual representations to parts of represented items. Thus, for example, a realistic sculpture of Abraham Lincoln is an iconic representation of Lincoln, and a photograph of Mao Zedong is an iconic representation of Mao. The point I wish to make here is that perceptual representations show some evidence of iconicity.

This follows from our earlier discussion of vectoral representations that consist of firing rates of individual cells. These vectors can be said to have parts. Moreover, the parts can be said to correspond to "parts" of the items that are represented. For example, parts of the vectors in the state space that represents faces correspond to features of faces, such as distance between eyes and fullness of lips. (These are the features in Churchland's radically simplified face space. The features in Chang and Tsao's more realistic model are more complex.) To be sure, features of faces are not parts of faces, except in an extended sense, but there is nonetheless an analogy here with paradigmatically iconic representations like sculptures and photographs.

There is also evidence of isomorphisms from parts to parts in a less metaphorical sense of "part." I have in mind Kosslyn's work on imagistic representations of a two-dimensional map (Kosslyn 1973, Kosslyn 1980). Kosslyn showed his subjects a map of a fictitious island with various landmarks, asked them to form images of the map, and then requested that they start at one landmark and scan across the image until they arrived at another. Kosslyn found that scanning times corresponded to distances between landmarks on the physical map. How best to interpret this result has been hotly contested, but Kosslyn's preferred interpretation is quite natural. On his view, the experiment shows that the mind can construct structured representations with parts over which a distance metric can be defined, and that distances between the parts of the representations correspond to distances on two-dimensional surfaces. We need not suppose, however, that a distance metric defined over neural representations requires that something with the physical shape of a map occurs anywhere in the human brain, Thus, as Michael Rescorla points out in this connection (Rescorla 2018), there is a highly abstract, undemanding notion of distance according to which distances need not be distances in physical space. According to this notion, a distance function is any

function D from pairs of points in a space to the real numbers such that, for any two points x, y, and z

$$D(x\,y) \geq 0$$
$$D(x,\ x) = 0$$
$$D(x,\ y) = D(y,\ x)$$
$$D(x,\ z) \leq D(x,\ y) + D(y,\ z).$$

As Rescorla remarks, "A metric space may be composed of any entities whatsoever" (2018, 40). Hence, if we were to interpret Kosslyn's images as iconic representations of a two-dimensional external map, we would not be committed to viewing the images as pictures of the map in anything close to a literal sense.

A second example of the iconic nature of imagery is afforded by an experiment by Ronald Finke and Steven Pinker (1983). Finke and Pinker showed their subjects screens containing four dots. The dots disappeared and were then followed by an arrow that might or might not point to a location where one of the dots had occurred. Subjects were asked to determine whether the arrow did indeed point to a location that had previously been occupied. Response times were measured. The key finding was that when an arrow did point to a formerly occupied location, the time required for a response was a linear function of the distance from the arrow to the target location. The natural interpretation of this finding is that the subjects superimposed an image of the screen with the arrow on an image of the screen with the four dots and traced a line from the tip of the arrow to the target. The longer the line that had to be traversed, the longer the time required. This interpretation is consistent with the subjects' testimony as to the procedure they had followed in arriving at an answer. Assuming it is correct, the experiment supports the hypothesis that the imagination uses representations with identifiable parts and a distance metric defined over the parts.

Both Kosslyn's experiment and the experiment by Finke and Pinker are concerned with the imagination. But we are principally interested in perception. Do the findings of these investigators have any implications concerning perceptual representations? Yes. As we noticed in the introduction to this chapter, there is strong evidence coming both from behavioral studies and from neuroscience indicating that the states and processes associated with the imagination are fundamentally akin to the states and processes associated with perception. If the imagination uses iconic representations, it is quite likely that perception does so as well.

Assuming that the iconicity hypothesis is correct, for at least some aspects of perceptual representation, we have identified another respect in which perceptual representations differ from conceptual representations; for it is questionable

whether there are ever isomorphisms between individual conceptual representations and their semantic values, and even if isomorphisms exist in some cases, the resulting iconicity is more circumscribed than it is in the case of perception. To be sure, as Wittgenstein in effect pointed out, if we think of Fregean propositions of the forms *a is F* and *a R b* as standing for structured entities with objects, properties, and relations as constituents, then propositions of the given forms can be said to be isomorphic to the entities they signify (Wittgenstein 1953). For example, on Wittgenstein's view, every Fregean proposition of the form *a R b* corresponds to a three-part structure consisting of the referent of *a*, the referent of *R*, and the referent of *b*. But there is no compelling reason to accept Wittgenstein's view that the states of affairs signified by such propositions are structured entities. Moreover, there are no relations among the parts of Fregean propositions that are counterparts of the distance and direction relations that we found to characterize the parts of imaginative and perceptual representations.

I will not expand on this fourth difference further here, for it is unclear how extensive the role played by iconicity in spatial representation might be. Where does iconic representation leave off and non-iconic representation begin? This is a difficult question that requires extended discussion.

One point is that some perceptual representation is pretty clearly not iconic. Consider, for example, a case in which a vector of firing rates represents a particular shade of orange. Suppose that the shade can be described as 40% yellow and 60% red. On the face of it, yellow and red aren't parts of the given shade of orange, in any sense of "part," so it would be quite implausible to maintain that this is a case of iconic representation. (In the case of face space, it makes sense to think of a vector as standing for a conjunction of facial features, but shades of orange are not conjunctions of other colors.)

5. A fifth difference is that perceptual representation is highly perspectival. I argue for this claim at length in Chapter 2. There is no need to repeat those arguments here. (Pautz also defends this view in Pautz 2021 (131).) I should say, however, that while actual perception is highly perspectival, this appears to be a contingent feature of perception rather than a necessary feature. To illustrate, Biederman argued that upper-level vision contains primitive non-perspectival representations of various shapes that he called "geons," and that these geons could be combined to generate representations of virtually all three-dimensional structures (Biederman 1987). Attempts to verify Biederman's conjecture experimentally ran into serious difficulties (Bülthoff and Edelman 1992, Tarr and Bülthoff 1995), but the claim that there is a perceptual system that makes use of geon-like representations seems perfectly coherent. Moreover, there is recent evidence suggesting that the earlier objections to Biederman may have to be qualified. It appears that in some cases humans can rapidly infer perspective-invariant representations of shapes from single perspectival representations (Baker and Kellman 2018).

The previous section made a case for there being two significant differences between conceptual representation and perceptual representation, and in this section we have added five more. We can summarize these findings by saying (i) that the expressive powers of conceptual representation are quite different than the expressive powers of perceptual representation, and (ii) that the encodings of information are quite different in the two cases—conclusions that help us to understand why Mother Nature went to the trouble of endowing us with both. Interestingly, however, these two representation schemes are alike in one extremely important respect. This is the topic of the next section.

V. Classes of Worlds as the Semantic Values of Perceptual States

It is a truism that perception guides us as we negotiate the physical environment and operate on it, making it possible for us to avoid trees, climb rocks, chip flints, and type on keyboards. It appears that representational theories of perception provide the best explanation of these and many other forms of guidance. We are able to avoid trees and climb rocks because perceptual states represent trees and rocks, and also their properties and relations, including their egocentric relations to ourselves. It is natural to conclude from this that we are able to negotiate and operate on the environment because we are endowed with representations that stand for circumstances in the environment—that is, arrangements of environmentally situated objects, properties, and relations. Since arrangements of this sort can be modeled by classes of possible worlds, we can express this conclusion by saying that the veridicality conditions for perceptual representations are such classes.

Before fully embracing the conclusion, however, we must consider two issues that might be sources of concern.

First, we should consider whether there is any reason to think that logical structure contributes to the individuation of the entities designated by perceptual representations. If so, it would be a mistake to think of those entities as classes of possible worlds, for, as we observed in Section I, logical structure has no bearing on the individuation of possible worlds as standardly understood. Now, we have found that perceptual representations do not involve close counterparts of logical concepts like conjunction and disjunction. We have also found, however, that perception involves operators that are remote counterparts of certain logical concepts—specifically, conjunction and existential quantification. Does the fact that these quasi-logical operators contribute to the individuation of perceptual representations have any tendency to show that the semantic values of those representations have structural elements corresponding to the operators? No, as can be seen by reflecting on the comparable case of Fregean propositions that contain

full-blown logical concepts. A Fregean proposition of the form *P and Q* is true just in case P is true and Q is also true. But exactly the same holds for Fregean propositions of the form *it is not the case that (either it is not the case that P or it is not the case that Q)*. Despite being radically different in point of logical structure, corresponding propositions of these two forms are true in exactly the same circumstances. That is, they have the same truth conditions. And of course there are infinitely many other Fregean propositions that share those truth conditions, involving increasingly baroque combinations of logical concepts. Accordingly, it is a mistake to suppose that the entities that serve as truth conditions for conjunctive Fregean propositions are individuated by a structural element corresponding to conceptual conjunction. If we were to suppose that, we would also have to suppose that they are individuated by infinitely many other combinations of structural elements. Similarly, Fregean propositions of the form *there exists an object that is F* have the same truth conditions as Fregean propositions of the form *it is not the case every object fails to be F*, despite diverging sharply from the latter propositions in point of logical form. This shows that the semantic values of existential propositions should not be thought to contain an element corresponding to the quantifier. The general moral here is that the semantic values of representations do not have elements or structural features corresponding to all of the items that contribute to the structure of the representations. It is enough that they contain objects that possess the properties that stand in the relations that the representations express. There is no argument from the logical or quasi-logical forms of representations to conclusions about the contents or structure of represented entities.

This brings us to a second possible concern about the view that the semantic values of perceptual representations are classes of worlds. We often characterize perceptual experiences using descriptions like "the man who is smiling" and complex demonstratives like "that smiling man." Thus, we might say that Josh sees the man who is smiling. That is, we use singular terms rather than sentences in describing perceptual states, even though the semantic values of singular terms are individuals rather than states of affairs. This practice can suggest that the primary objects of perceptual awareness are individuals of some kind. Susanna Schellenberg seems to have embraced this idea, and there are passages in the writings of Tyler Burge which suggest that he finds it at least tempting. (For example, "Perception is essentially, at every point, context-bound singling-out of particulars" (Burge 2010, 542).) But the idea seems quite wrong to me. When we characterize perceptual experiences in terms of singular descriptions or complex demonstratives, we are prescinding from representational contents that are best seen as states of affairs. Consider the fact that when I see Adam as smiling, I also see him as a constituent of a state of affairs containing a number of other objects that are bound together by a complex web of relations. In a particular case, I may see him as smiling, but also as seated at his desk, looking at a computer screen,

and surrounded by bookshelves. There is no way that a complex situation of this sort can be analyzed purely in terms of individuals and monadic properties of individuals. It is impossible to analyze relational facts in terms of monadic facts.

We earlier took note of a strong reason for thinking that perceptual representations stand for classes of worlds, and we have just now found that two worries about this thesis are groundless. In view of all this, it seems reasonable to embrace the thesis, at least tentatively. Or, to be precise, it seems reasonable to embrace the thesis that perceptual representations stand for classes of centered possible worlds. Since this is also true of beliefs (in the sense that classes of worlds are one of the two types of objects of belief), we have found an important commonality. (The semantic values of beliefs should also be thought of as classes of centered worlds because beliefs can contain indexical and demonstrative concepts whose significance depends on the current point of view of the believer (Lewis 1979).) Among other things, this commonality helps us to understand how perceptual experiences can play a large role in ostensive definitions and other ways of fixing the referential contents of concepts, and how experiences can play a causal role in producing beliefs with a full share of external content.

VI. Conclusion

Although cognition has many dimensions and makes use of representational phenomena of a number of different kinds, including maps and images, it relies heavily on beliefs and other propositional attitudes. These states lack the fine-grained expressive powers of perceptual states, in the sense that they do not easily register highly determinate values of qualitative properties or highly intricate spatial structures. They are, however, essential in registering less particularized subject matter. Attitudes also have the following features: they are constructed from concepts; they are largely independent of control by current stimuli; they are individuated by logical forms; they figure in reasoning that is sensitive to these forms; they are neither analog nor iconic; and they are not essentially perspectival (even though some of their number contain indexical elements that are keyed to perspectives). We have found that perceptual states differ from attitudes in all of these respects. The only significant point of similarity is that attitudes and perceptual states have semantic values of the same kind. We have also noticed, however, that similarity in this respect is important: it makes it possible for cognition and perception to cooperate normatively, informationally, and causally.

9
The Epistemic Role of Perception

I. Introduction

It is widely held that an agent X knows that a proposition P is true just in case the following four conditions are satisfied: (i) X believes that P; (ii) it is true that P; (iii) X is justified in believing that P; and (iv) X isn't Gettierized with respect to P. X is Gettierized with respect to P if conditions (i)–(iii) are fulfilled, but there are abnormalities in the context that make it a lucky accident that X's belief that P is true. Here is a famous example, due to Keith Lehrer (Lehrer 1979), of an agent who is Gettierized:

> You see Nogot driving a Ford on a regular basis. Nogot tells you he owns the Ford. You've never seen Havit drive at all. He walks everywhere or takes public transportation. There are three people who work in the office—you, Nogot and Havit. You don't own a Ford. You reason like this: Nogot owns a Ford, and Nogot works in the office, and so someone who works in the office owns a Ford. It turns out that Nogot doesn't own the Ford he's been driving at all. Moreover, Havit does own a Ford, but it's an old classic car he never drives. You have a justified true belief that someone in the office owns a Ford, but do you know someone in the office owns a Ford?
>
> (I quote from Goldman and McGrath 2015, 39.)

The theory of knowledge based on (i)–(iv) is supported by an extremely broad range of examples. When one considers actual and possible agents who count intuitively as possessing knowledge, one always finds that they satisfy the four conditions; and when one considers cases that fail to satisfy one or more of the conditions, one is strongly disinclined to classify them as examples of knowledge. For this reason, I will assume here that the theory is correct.

There are two important ways in which perception contributes to knowledge: perceptual states can cause beliefs (condition (i) in the foregoing theory of knowledge) and can also contribute to epistemic justification (condition (iii)). I will be concerned with both of these contributions in the present chapter, though I will have more to say about the latter contributions than the former. Perception has no bearing on the truth component of the foregoing theory (condition (ii)), and very little bearing on the Gettier component (condition (iv)). The issue of whether

Perceptual Experience. Christopher S. Hill, Oxford University Press. © Christopher S. Hill 2022.
DOI: 10.1093/oso/9780192867766.003.0009

a belief is true depends entirely on whether the proposition that serves as the object of the belief stands in an appropriate correspondence relation to the world. What the believer is perceiving is irrelevant to that issue. Further, as the Gettier example involving Nogot and Havit illustrates, the bearing of perception on the fourth, Gettier condition is uninteresting. To be sure, we can say that you would not have been Gettierized if you had perceived the vintage Ford in Havit's garage, but saying that is hardly illuminating. It is a special case of the generalization that failure to perceive relevant facts can prevent us from acquiring knowledge. This generalization is a virtual tautology, for its occurrence of "relevant" clearly means "knowledge-relevant." We gain virtually nothing from knowing that the pursuit of knowledge can be hindered by failing to perceive facts that are germane to the pursuit's success.

Now the type of justification that figures in knowledge is just one of several different types of doxastic justification. To distinguish it from others, I will refer to it as *epistemic justification*, a label that is warranted by the *OED* and other dictionaries. Thus, lexicographers speak with one voice about the meaning of "epistemic": they all take it to mean something like "of or relating to knowledge." Adopting this usage, I can describe the main thrust of the present chapter as an effort to delineate the role that perception plays in epistemic justification. The focus will generally be on the part of the role that belongs to perceptual experience.

To appreciate that there are other types of doxastic justification, note that there is a sense in which agents might be justified in holding certain religious beliefs. Suppose that B is a religious belief that cannot be decided on the basis of evidence, and consider an agent who holds B because it provides a sense of meaning or teleological depth—a reason for living that the agent would not otherwise have. In considering such a case, it seems natural to share William James's opinion that the agent is justified in holding the belief, though, unlike James, one would probably want to be quite explicit that the form of justification is purely practical and has no bearing on matters epistemic (James 1960). I will consider several forms of non-epistemic doxastic justification towards the end of the chapter—all quite different than this sort of practical or pragmatic justification—and will consider what bearing perception might have on them.

There are two main ways of thinking about epistemic justification. *Externalist* theories maintain that whether a belief is epistemically justified depends on conditions that are external to the believer, such as whether the belief is produced by a faculty that has a strong tendency to produce beliefs that are true. (Here and elsewhere, I will be assuming that truth is a relation between beliefs and external states of affairs.) *Internalist* theories take the opposite position. According to them, epistemic justification depends entirely on factors that are internal to the believer. *Access internalism* maintains that all of these factors are immediately accessible to the believer—that the believer can grasp them fully from the first-person perspective. The form of internalism known as *mentalism* does not impose

an accessibility requirement on epistemic justification. Rather, it just asserts that all of the factors on which justification depends must be internal in the sense of being mental states or mental processes of the believer. It allows that many of these states and processes are beyond the reach of introspection.

I will be principally concerned to show that the role of perception in epistemic justification is best described by externalist theories. They make sense of our intuitions about the role of perception and also possess other desiderata. In contrast, access internalism and mentalism fail to do justice to the *epistemic* nature of the form of justification that is our target. Before taking up this business, however, it is necessary to introduce an additional concept, root out an error, and make some foundational claims.

II. Experiential Judgments and Experiential *Of*-ness

We know from Chapter 2 that the contents of perceptual experiences are concerned with external objects and the appearance properties of such objects, and we know from Chapter 8 that the contents of experiences are autonomous in the sense that they owe nothing to concepts, however closely concepts may be bound to them by causal ties, and however difficult it may be to distinguish their contents from conceptual contents introspectively. On the other hand, beliefs of all kinds, including the occurrent beliefs that we call *judgments*, are constitutively conceptual in nature. They have concepts as their building blocks.

A perceptual judgment J counts as *experiential* just in case (i) it is caused directly by a perceptual experience E and (ii) its content coincides with the autonomous content of E. Given the observations of the preceding paragraph, it follows that J is constructed from concepts whose contents capture parts of the autonomous content of E. More specifically, J is constructed from a concept that picks out an external object and concepts that stand for appearance properties of that object. (I am compressing a bit here. In view of Chapter 4, we know that the contents of perceptual experiences can be fully expressed by existential propositions. Their contents don't involve specific objects. But in order to simplify the exposition, I will often prescind from this fact in the present chapter.)

This definition permits us to formulate and evaluate a principle about the epistemic role of perceptual experience which has some currency in the literature, but which I believe to be profoundly mistaken. Here it is:

(GT) The occurrence of a perceptual experience is such as to logically guarantee the truth of the experiential judgment that it causes.

The idea that the truth makers for experiential judgments are somehow built into the corresponding perceptual experiences seems to derive principally from a

sense that the truth makers are *presented* in experience—a sense that they have a *presentational immediacy* that can only be explained by supposing that the objects of experience are constituents of experience. On this view of presentational immediacy, there are no other constituents of experience that could sully our epistemic grasp of its objects. Perceptual experience doesn't involve perceptual representations, conceptualizations, or interpretations of any other kind. Presentational immediacy is also thought to preclude theories of experience that attribute an internal structure to the relation *experience of*, for structure can introduce distortions. Parts of the external world are *given* in experience, and we must understand experience in a way that allows for that givenness. We can only do this by supposing that perceptual experience *incorporates* external phenomena. (For an extended defense of (GT) that is based on presentational immediacy, see Beck 2021.)

What we are seeing here is an account of presentational immediacy which carries an epistemological benefit (support for (GT)), but which commits us to a very substantial metaphysical claim about the relationship between perceptual experience and its objects. I suggest that the metaphysical claim is untenable. If we were to accept the idea that perceptual experience incorporates external phenomena, we would be unable to make sense of illusory and hallucinatory experiences. Advocates of the incorporationist view try to meet this objection by maintaining that illusory and hallucinatory experiences belong to a different natural kind than the corresponding veridical experiences. Veridical and nonveridical experiences lack a common nature, except for sharing the uninteresting disjunctive property *either a veridical experience or a hallucination*. In a phrase, the advocates maintain that perceptual experience is *disjunctive* rather than unitary (Byrne and Logue 2008). There is much that is wrong with this disjunctivist view, but here I will just point out that its adherents must shoulder a very large burden of proof. Neither introspection nor neuroscience can discern any difference between vivid illusions and hallucinations (such as those that occur in patients with the Charles Bonnet syndrome; see Chapter 4) and veridical experiences, so there is an appearance of kind-identity that disjunctivists must explain away. The last time I looked they had not made much progress on this task.

Anyway, disjunctivism aside, the support that presentational immediacy provides for (GT) is minimal. This is because (i) the doctrine of immediacy is evidently grounded in phenomenology, which creates an impression of immediacy, and (ii) this phenomenology can be interpreted in more than one way. In fact, the phenomenology of immediacy can easily be explained by representationalism, for, when it is properly understood, representationalism implies that experiential awareness is immediate and direct. Thus, according to representationalism, the representations that are constitutive of perceptual experience stand only for external phenomena. There are no mediating entities that figure in their representational contents.

We have been considering a reason why perceptual experience might be thought to boost our epistemic stature, by guaranteeing the truth of the experiential judgments that it occasions. There is also a reason for thinking that it might detract from our stature. To appreciate this reason, we need to take note of the fact that the expression "experience of" is ambiguous. In one sense of the expression, claims of the form "X is enjoying a perceptual experience of A" entails that A exists, but in another sense it does not. Thus, for example, it can be true that X is having a perceptual experience of a unicorn. We can express this ambiguity by saying that "experience of" has both an *extensional* meaning and an *intensional* meaning. (To say that a term has an intensional meaning is just to say, first, that it isn't existence entailing, and, second, that truth values aren't preserved by replacing terms within its scope with coreferential terms. For example, "fears" has an intensional meaning: I can fear Cerberus, even though Cerberus doesn't exist, and the fact that I fear Cerberus doesn't entail that I fear Aphrodite, even though "Cerberus" has the same reference as "Aphrodite" in that they both fail to refer. A term has an extensional meaning if its meaning isn't intensional.)

The complex expression "experience *as of*" is often used to emphasize that the intensional sense of "experience of" is in play.

Now, according to certain philosophers influenced by Franz Brentano, when we say that an experience is *of* an object using the intensional sense of "experience of," part of what we are claiming is that the experience is directed on a special kind of an object—an object that doesn't exist but nonetheless has a certain kind of being or reality (Brentano 1973, Parsons 1980, Meinong 1981). These special quasi-entities are said to be *intentional objects*. This is of epistemological interest because the view creates difficulties for the idea that perceptual experience puts us directly in touch with actual external objects. It implies that perceptual awareness of actual external objects is always mediated by experience of intentional objects, and is therefore dependent on inference. It is like the doctrine of sense data in that respect. Like sense data, intentional objects threaten to screen off actually existing objects, making it problematic that we are veridically aware of the latter.

Fortunately, the difficulties in this area are illusions. An intensional claim to the effect that an experience is *of* an object can be understood as a claim about the representational content of the experience, and claims about representational contents need not be understood as claims about relations to quasi-entities. As we saw in Chapter 4, they can be understood instead to be claims to the effect that representing states represent certain properties as instantiated. For example, to say that Macbeth's hallucinatory state represents a dagger is just to say that the state represents a certain cluster of properties as instantiated. There is no need to suppose that the first claim entails that Macbeth's state stands in a relation to a quasi-entity that possesses those properties. Equally, to say that a hallucinatory

state represents Pegasus is just to say that the state represents certain properties that legend ascribes to Pegasus as instantiated (and perhaps also that the hallucination primes the concept of Pegasus in the hallucinating agent).

Another preliminary point about experiential *of*-ness. I suggest that what experiences can be *of* is quite limited. As we saw in Chapters 2 and 3, perceptual experience can be said to be *of* appearances presented by objects. Further, as we saw in Chapter 1, it makes sense to suppose that there is more than one type of representation relation, and to suppose that a state can simultaneously stand in two or more such relations. In view of this, we can perhaps also say perceptual experiences can represent, and therefore be *of*, various higher-level phenomena such as numerosities, causal relations, and human faces. What I would like to consider now, however, is the fact that we sometimes say that we have perceptual experiences of *very* high-level phenomena such as pine trees, sheep, tables, vases crafted in the Ming era, Russian oil tankers, and senators. What could claims of this sort mean? I suggest that in such cases we aren't speaking only of the contents of perceptual experiences, but rather of the contents of complex representations that have both experiential and conceptual components. More particularly, I suggest that the truth makers for such claims are facts involving causal relations between perceptual experiences and higher-level representations. When a perceptual experience occurs, it immediately triggers post-experiential representations in the upper levels of the relevant perceptual system, and its influence then rapidly spreads to parts of the brain that house conceptual representations. Once these representations have been activated, they can combine with purely experiential representations to guide cognition and action.

According to this picture, the truth condition of a claim like "I see a Russian oil tanker" is determined in part by the content of a current perceptual experience and in part by the contents of concepts that the experience has activated. It isn't determined by the content of the experience alone. By the same token, according to the picture, it is a mistake to follow philosophers like John McDowell (McDowell 1996) in speaking of the conceptual content *of an experience*. Insofar as it is appropriate to speak of conceptual content in relation to experience, one could only be referring to the contents of complex mental states involving both experiences and concepts.

A related point is that it is misleading to speak of perceptual experiences as contained in "conceptual envelopes," even if one acknowledges that this is just a metaphor (Block 2022). The relationship between experiences and concepts is a causal relationship, not a relationship of containment. More particularly, experiences are not sealed off from cognition in the way that the envelope metaphor suggests. To be sure, the activation of concepts by perceptual experiences is normally automatic and immediate, permitting no purchase on how things stand in the perceived world independently of conceptualization. But what happens in

normal cases is only part of the story. There is also such a thing as patient, careful introspective access to experiences, and when we are enjoying such access, we can distinguish between experiential content and conceptual content. To elaborate: when we do what is called introspectively examining an experience, our first step is to invest the experience with attention so we can get a better fix on what the experience is of. We then describe what it is of using concepts or words. (See the discussion of introspection in Chapter 7, Section IX.) Here the choice of concepts or words is much slower and more deliberate. And it is reflective: we recognize that there is such a thing as being wrong in our conceptualizations, and we may correct our initial choices. ("No, my first attempt to describe that flower was wrong. Its color is more cerise than French rose.") In such cases, there is no sense in which we are sealed off from our experiences or their objects by reflexive conceptualization, though of course one is restricted in one's ability to describe the experienced world by the limitations of one's conceptual vocabulary.

In this section I have been trying to get a better fix on the features of perceptual experience that are germane to its epistemic potential. I will assume going forward that presentational immediacy has limited epistemological significance, that experience doesn't create a screen of intentional objects separating us from actual external objects, that we should not infer from ordinary usage that experience can represent such high-level properties as *oil tanker* and *senator*, and that we aren't sealed off from the experienced world by conceptualization in a way that idealists might find congenial.

III. A Brief for Process Reliabilism

Process reliabilism is a theory of epistemic justification due to Alvin Goldman (Goldman 1979). In a recent formulation of his view (Goldman and McGrath 2015, 34), Goldman summarizes it as follows:

> Belief B is justified if and only if B is produced by a reliable belief-forming process, i.e., a process that has a tendency to generate (belief) outputs with a high percentage of truths.

The focus, then, is on psychological processes that produce beliefs, and on the reliability of such processes. The main goal of the present section is to show that the account of epistemic justification that this view offers is highly attractive. In the Section VI I will consider some aspects of the justificatory roles that reliabilism assigns to perceptual experiences and perceptual states generally.

Now there are three different types of belief-forming process. First, there are processes that take items other than beliefs—specifically, perceptual experiences

and perceptual memories—as inputs and yield beliefs as outputs. Second, there are processes that take both beliefs and items other than beliefs as inputs and yield beliefs as outputs. The other items relevant here are perceptual experiences and perceptual memories. And, third, there are processes that take beliefs as inputs and yield beliefs as outputs. Processes of all three of these types can be described as inferential, though many of them have a kind of automaticity and inflexibility that higher-level, conscious inferences lack. Processes of the first type are said to be *belief-independent* and processes of the second and third types are said to be *belief-dependent*.

Because of these differences among the types of process that can produce beliefs, an adequate reliabilist account of epistemic justification has to be more complex than the highly compact formulation given two paragraphs back. An account must explain justification for beliefs produced by belief-independent processes in a different way than justification for beliefs produced by belief-dependent processes. Here is the expanded account that Goldman presented in his first paper on reliabilism (Goldman 1979):

> If S's belief in P at [time] T results from a belief-independent process that is reliable, then S's belief in P at T is justified.
>
> If S's belief in P at T results from a belief-dependent process that is *conditionally* reliable, and if the beliefs on which this process operates in producing S's belief in *P at T are themselves justified, then* S*'s belief in P at T is justified*.

This formulation has continued to enjoy preferential status, both in Goldman's own work and in that of other reliabilists, across the years. (A belief-forming process is said to be conditionally reliable if its outputs have a high proportion of true beliefs in cases in which the input beliefs are true.)

Since reliability is defined in terms of truth, and reliabilism assumes that truth consists in correspondence with extramental fact (in the case of propositions that are not concerned with the mind), reliabilism is a paradigmatic form of externalism.

Many philosophers find reliabilism to be a highly appealing—even a highly compelling—account of epistemic justification. Its popularity is due to several virtues. The first virtue is that it fits our intuitions about concrete cases of justification. We tend to think of beliefs produced by reliable processes, such as inference from perception, as justified, and of beliefs produced by unreliable processes, such as wishful thinking, as unjustified. The second virtue is that reliabilism has an easy and fully satisfying answer to the question, "Why should I care about having beliefs that are epistemically justified?" By definition, beliefs formed by reliable processes are likely to be true, and we all know that actions based on true beliefs are more likely to succeed than actions based on beliefs that are false.

This is no less true of purely intellectual actions than of actions aimed at an external goal. Third, reliabilism implies that the second and third conditions cited in the standard theory of knowledge are deeply unified, for it implies that satisfaction of the third condition entails that the relevant belief is likely to be true. It therefore helps us to understand the motivation for the concept of knowledge—why we use a single concept to keep track simultaneously of just the factors that figure in the forementioned four conditions. The fourth virtue is that reliabilism excludes certain forms of *epistemic luck*. Intuitively, a belief cannot count as knowledge if it is just a lucky break that the belief is true. For example, we would not classify a belief produced by guessing as knowledge even if by accident it happened to correspond to a fact. But if a belief is produced by a process that yields true beliefs in a large percentage of cases, it can hardly be luck if it turns out to be veridical. This fourth virtue is closely associated with a fifth: reliabilism unifies the factors cited in the third and fourth conditions in the standard theory of knowledge. As we have just seen, reliabilism implies that the third condition is an anti-luck condition. But reflection on Gettier cases shows that they all involve beliefs that are true owing to a kind of luck. Versions of the fourth condition can all be seen as attempts to rule out luck of this sort. Hence, if reliabilism is true, the third and fourth conditions in the definition of knowledge are addressing different aspects of the same issue.

I believe that internalist accounts of epistemic justification lack the first of these virtues, but making that case requires extended treatment. I will defer that discussion to Section VII. In this section I will just point out that internalist theories necessarily lack the remaining four virtues.

This is easily accomplished. (1) As we saw, a reliabilist can explain why it is important to us that our beliefs be epistemically justified by pointing out that beliefs that are likely to be true are also likely to help us achieve success in pursuing our goals. Internalism would be hard put to come up with a virtue that is more appealing than a close link with goal fulfilment. To be sure, some of the features of beliefs that internalist theories cite in explaining epistemic justification are desirable. An example is coherence among the whole system of one's beliefs. But purely internal features like coherence pale in significance in relation to probability of truth. They carry no promise that pursuit of goals involving external objects is likely to succeed. (2) The third virtue is that reliabilism unifies the second and third conditions in the standard theory of knowledge: according to reliabilism, fulfilment of the third condition makes it likely that the second condition is also fulfilled. This establishes that knowledge has an important kind of internal unity. It is clearly impossible for an internalist theory to match this result, for no purely internal feature of a belief can by itself increase the chance that the belief is true. (3) The fourth virtue is that by explaining the third condition in the standard theory in terms of reliability, reliabilism rules out the kind of epistemic luck

that occurs when one happens to guess correctly. Since this sort of luck is defined in terms of truth, internalist theories have no bearing on it. (4) Like the third virtue, the fifth involves unity. More specifically, in interpreting the third condition in such a way that it rules out a certain sort of epistemic luck, reliabilism reveals a close relationship between the third condition and the fourth, which is also concerned to block epistemic luck, though luck of a different kind. According to reliabilism, as we have just seen, the third condition rules out the kind of luck that occurs when one arrives at a true belief via an untrustworthy process like guessing. The fourth condition rules out the kind of luck that occurs when one arrives at a true belief despite being in an epistemically hostile context (that is, a context with hidden features that would normally prevent a belief from being true). Thus, reliabilism brings to light a second way in which the concept of knowledge is unified. Internalism cannot match this result. Both forms of epistemic luck are defined in terms of probability of truth. Accordingly, internalism has nothing to say about them.

It is worth emphasizing the fact that internalism lacks the unifying power of reliabilism. Even if it had no other failings, it would be in quite serious trouble for this reason. Thus, it would be faced with an important question that it could not answer—the question of why we should possess, care about, and frequently deploy a concept with four elements that it treats as disparate. There must be some sort of glue that holds the elements together, and makes it natural for us to want to keep track of them simultaneously. But internalists are ex officio excluded from explaining that glue.

I have been maintaining that reliabilism has four virtues that are not shared by internalist theories of epistemic justification. I will now add a fifth. Let us begin by asking, "Why is it useful to have a concept of knowledge that involves a justification condition and a Gettier condition as well as a truth condition?" Why isn't it sufficient to attribute beliefs and to assert that they are true? The reason, I suggest, is that when we attribute knowledge that P to a subject, in addition to claiming that the subject believes that P and that the belief is true, we are indicating that there is *evidence* that the subject's belief is true. This evidence has two components. First, we are citing the subject's track record with respect to beliefs formed by the same process as the current belief that P. Given that the subject has been reliable about such matters in the past, this is evidence for the truth of the agent's current belief. And, second, we are calling attention to the subject's current epistemic environment, and claiming that it is normal in ways that make the subject's track record relevant. This is also evidence for truth. In effect, then, when we ascribe knowledge to an agent, thereby claiming implicitly that the agent's belief is produced by a reliable process and that the agent is not in a Gettier situation, we are claiming there is evidence that the agent's belief is true. Note that this evidence *might* only be known to the *attributors* of knowledge. The *attributee* might not know

that the relevant process is sufficiently reliable to count as justification-conferring, and/or might not know that there are no Gettier elements in the relevant context.

What we are seeing here is another way in which the concept of knowledge is unified, assuming that reliabilism provides a correct account of epistemic justification. On that assumption, we can see that attributions of knowledge both ascribe truth to a belief and point to evidence which supports that ascription. There is cooperation among the various conditions that define knowledge.

The thing to note is that this argument presupposes a reliabilist construal of epistemic justification. If justification consisted in satisfying a purely internalist condition, then the fact that the relevant agent has justification for believing that P would not ipso facto provide strong evidence for the truth of P.

I think we can see that we have the purpose of citing evidence that supports ascriptions of truth no less when we attribute knowledge to ourselves than when we attribute it to others. In both cases, we are claiming that a belief is true and also calling attention to the fact that there is indirect evidence of certain kinds that it is true.

IV. Skepticism

It is another advantage of reliabilism that it provides a persuasive answer to radical skepticism—that is, the form of skepticism that purports to call into question *all* of our empirical knowledge of the world (Hill 1996).

Here is how a radical skeptic reasons: "It could be the case that you are the victim of a Cartesian demon, or that you are a brain in a vat whose perceptual experiences are caused by events inside a supercomputer. In saying that these hypotheses could be true, I mean that you are not in an epistemic position to rule them out: you are not epistemically justified in rejecting them. But they contradict all of your empirical beliefs about the external world. It follows that you are not epistemically justified in holding those beliefs."

The reliabilist can respond as follows: "You claim that we are not epistemically justified in rejecting your skeptical hypotheses, and that we lack epistemic justification for holding our empirical beliefs. But, according to reliabilism, a belief counts as epistemically justified just in case it is produced by a reliable process. Hence, to establish your claims, you must show that the processes that produce our empirical beliefs are in fact unreliable. That is, you must establish that certain claims about the external world are true. But, by your own lights, it is impossible to do this! So if reliabilism is true, your position is untenable."

Of course, it is also true that the reliabilist cannot give the skeptic an argument that our perceptual processes *are* reliable that the skeptic won't regard as question begging; but that just means that there is a standoff at the second level between

skeptics and reliabilists. It doesn't affect the fact that the reliabilist has succeeded in blocking the skeptic's claim that our first-level perceptual beliefs are unjustified.

To expand on the standoff: If reliabilism is true, the skeptic can't prove that our empirical beliefs are not epistemically justified while continuing to maintain that empirical knowledge is impossible; but, also, the reliabilist can't prove that we *are* justified in holding our empirical beliefs without making assumptions about reliability, and therefore about the external world, that the skeptic will regard as question begging. Neither of these opponents is in a position to convince the other. But acknowledging this second-order impasse is perfectly compatible with recognizing that, according to reliabilism, the skeptic has failed to show that our empirical beliefs are unjustified.

I emphasize that these considerations apply only to *radical* skepticism. This very robust form of skepticism will also be the target of my remarks in Section VII. To be clear about the intended scope of these discussions, I should distinguish radical skepticism from a currently popular form of skepticism according to which it follows from *empirically certified* premises that it is *in fact likely* that our lives are not what we unreflectively take them to be, but rather simulations that are produced by sequences of events inside computers (Bostrom 2003a and 2003b, Chalmers 2022). The foregoing remarks are intended to apply only to radical skepticism, as are the remarks about skepticism in Section VII. I would, however, like to pause briefly to consider this second form of skepticism. This will be something of a detour, but there would be little value to answering radical skepticism if an alternative, more solid form of skepticism remained in the field.

I will refer to the second form as *simulation skepticism*. Advocates such as Bostrom and Chalmers try to support it with versions of the following *simulation argument*.

First premise: There are trillions of planets in the universe I inhabit.

Second premise: Because the number is so vast, it is highly likely that there are a great many planets in this universe on which creatures have evolved who have achieved a level of technology that would enable them to produce simulated lives.

Third premise: Since we are still dealing with very large numbers, it is highly likely that the inhabitants of many such planets have been motivated to produce many simulated lives.

Fourth premise: The number of simulated lives under consideration in the third premise is so large as to make it likely that there are vastly more simulations of lives than real lives.

Conclusion: It is likely that my life is a simulated life.

The argument presupposes the notion of a simulated life. Briefly, a simulated life is a life in which all of the conscious experiences of an agent are due to a

computer running a complex program. Thus, for example, if my life is a simulated life, when I have experiences as of hiking in a forest, or as of typing on a computer keyboard, my experiences are caused by patterns of activity in the computing units of a machine of some kind. There are two main kinds of simulated life. One kind involves a real human being or brain all of whose experiences are simulations. This is the situation of the character Neo in the movie *The Matrix* (until he takes the red pill). The other kind involves a simulated subject with simulated experiences.

I find the simulation argument puzzling: I don't see where the probability estimates in the premises are coming from. As a statistician friend, John Lamperti, has pointed out, the second premise appears to overlook the fact that we humans are currently on the verge of destroying most of the life on the planet we inhabit. Perhaps it is likely that creatures with sufficient ambition and tunnel vision to build a civilization like ours will have a strong tendency to destroy the natural resources that are required to sustain that level of existence along the way. Further, consider the third premise, which asserts that biological creatures would very likely be strongly motivated to create numerous simulated lives with simulated subjects, if they were sufficiently advanced to be capable of doing so. This seems highly questionable to me. Biological agents generally need to have personal benefits in view in order to be motivated to undertake extremely complex projects. Their resources of energy are too precious to be squandered on unrewarding ventures. Now, there would no doubt be plenty of incentive to produce simulated life episodes for actual people. Who wouldn't want to spend a number of hours experiencing a Caribbean island in the dead of a northern winter, and who wouldn't be willing to pay for such an opportunity? But what is the advantage in simulating the complete life histories of simulated people? Who would pay for such *jeux d'esprit*? Also, it seems likely that most simulations would have telltale glitches.

My main concern here, however, is not to question the premises, but to point out a certain tension or opposition between the premises and the conclusion. The premises are all empirical claims about objects in the physical world around me. They are claims about planets, people, computers, programs, and motivation. They also presuppose many additional empirical claims about other physical objects—stars, galaxies, telescopes, telescope images, human bodies, human behaviors, and so on. Thus, it would be impossible to know that the premises of the simulation argument are true without knowing many other facts about, say, stars. Now the premises and their presuppositions distinguish sharply between things of the sort I have been mentioning—planets, stars, people, etc.—and the simulations that are allegedly occurring on other planets and allegedly will occur on Earth in the future. Thus, I could not know the first premise of the argument unless I knew about stars other than my own from having seen them from locations on Earth, and from combining those visual observations with facts about the total number of stars that I have acquired by interacting with fellow Earthlings.

But these things imply that I live on Earth and am therefore entirely independent of simulations that may be occurring on those other planets. They imply I stand at an immense distance from all of those simulations. In view of these considerations, we can see that knowledge of the premises presupposes that I am independent of the simulated lives to which the premises refer. And from this it follows that if I know that all of the premises are true, I cannot also know that the conclusion is true. This means that the argument fails. In a successful argument, knowledge of the premises provides one with knowledge of the conclusion, or at least a basis for obtaining knowledge of the conclusion.

I add that I could not know that there may someday be total simulations on Earth unless I had perceived contemporary computers and virtual reality equipment and had extrapolated from them. That is, I could not know about possible future simulations on Earth unless I had observed contemporary devices from the outside, and had drawn certain inferences about the future from them. This implies that I know that I stand at a considerable temporal distance from any total simulations that may appear on Earth.

An advocate of the simulation argument might at this point object as follows: "It still could be true that you are living inside a simulation! When you take yourself to be observing stars, telescopes, other people, and so on, it could be that your brain is just being stimulated by computer states, or even that you are yourself a series of computer states that are being activated by other computer states. If this were the case, the truth conditions for both the premises and the conclusion would just involve computer states of various kinds."

The operative word in this riposte is "could." Unlike the simulation argument itself, which purports to deliver a claim about the actual probability of a certain fact obtaining in the actual universe, the reply merely claims that we are not epistemically justified in ruling a certain skeptical hypothesis out. For this reason, unlike the simulation argument itself, the reply has the same form as the traditional arguments for radical skepticism. But this means that the reply has already been answered above. As we observed, process reliabilism shows that radical skepticism is untenable.

V. Answers to Objections

So far so good, but there are several unresolved issues about the proper formulation of reliabilism. To illustrate, there is the question of how to classify belief-forming processes in ways that are most relevant to epistemic justification. A given belief-forming process can always be classified in a number of different ways. For example, the process that produces a given belief about the shape of an object might be classified as vision, as spatial vision (as opposed to color vision), as attentive

spatial vision, and as attentive spatial vision qua distorted by mescaline. The reliability of the process obviously varies with the category in terms of which the process is classified. Which of the categories is most relevant to the epistemic status of the belief? This *generality problem* is challenging: the solutions that first occur to one turn out to be wrong. Fortunately, however, the generality problem seems to be yielding to assaults by reliabilists (Lyons 2019). I will not review this dialectic here, referring the reader to Lyons 2019 instead, nor will I take up any other issues about the proper formulation of reliabilism. (For discussion of these other issues, see Weisberg 2012, Goldman and Beddor 2015, Goldman and McGrath 2015, and Graham and Lyons 2021.) Instead I will try to bring out the strengths of the view by responding to two concerns about it that are motivated by internalism.

The first of the objections is the *new evil demon problem* (Cohen 1984, Pollock 1984). Here is a representative account of it:

> Descartes entertained the possibility of an evil demon that systematically deceived him, causing even his perceptual experiences to mislead him about his surroundings. Contemporary critics of reliabilism use the evil-demon hypothesis to pose a counterexample to reliabilism. Since the character in the evil-demon world is systematically deceived by his perceptual experiences, his perceptual belief-forming processes must be unreliable. Hence, according to reliabilism (it seems), he isn't justified in believing their outputs. But, these critics protest, such beliefs should intuitively be classified as justified, presumably because agents in an evil demon world have the same experiences as people in the actual, normal world. Ostensibly, this is a major flaw in reliabilism.
>
> (Goldman and McGrath 2015, 39)

Those who advance this objection evidently have the intuition that the character in the demon world is *epistemically* justified in trusting beliefs that are produced by such standard processes as perception and memory. Assuming that there is such an intuition, there must be a source of epistemic justification other than the reliability of belief-forming processes, for the processes in the example are altogether unreliable. Advocates of the objection conclude that epistemic justification must be internalist in character, or must at least have a substantial internalist component. (See Smithies 2019, 197–203, for a more expansive presentation of the argument.)

Although this objection has been widely influential, it yields to a fairly simple response. If the demon world seems to pose a problem for reliabilism, it is because there is a tendency to confuse epistemic justification with prudential or practical justification. Rather than allowing that the character in the demon world, whom I will call *Carol*, is epistemically justified in holding her beliefs about the external

world, a reliabilist can plausibly maintain instead that she is entitled to take her beliefs at face value because she has normally been rewarded for doing so in the past. Thus, for example, when it has seemed to Carol that there is appetizing food within reach, and on the basis of that impression she has undertaken what seemed to be the action of reaching for food, she has been rewarded by a pleasant taste and a reduction of hunger pangs. Because this pattern has almost always held in the past, Carol is fully entitled to trust that it will continue in the future. But her entitlement doesn't derive from the truth of her past beliefs. All of those beliefs have been false. Rather it results from the fact that her past beliefs have been *rewarded*. Induction combines with past experiences of reward to provide her with a prudential or pragmatic justification for continuing to form beliefs like ours by processes like the ones we use. (Goldman offers a quite different reply to the new demon world objection, which I find equally convincing; the objection is doubly flawed (Goldman and Beddor 2015, Goldman and McGrath 2015).)

The second objection to reliabilism starts with the observation that a theory of justification must cite a procedure for determining when a belief is justified. Applying this observation to reliabilism, we get the result that it must be possible for an agent, A, to determine which of A's belief-forming processes are reliable. Now to identify a process P as reliable, it seems that A must conduct an inquiry that runs as follows: "Hmmm, let's see. Belief B1 was produced by process P and B1 is true; . . . ; Bn was produced by P and Bn is true; so, by induction, P is reliable." But here we must ask, how does A determine that Bi is true? It might seem that A can only determine that Bi is true by using the very process P whose reliability A is trying to assess. But, if so, then any attempt to establish the reliability of P will be vitiated by circularity. Or so it might seem.

This line of thought seems to favor internalism because the crucial step is the observation that if reliabilism is correct, then determining whether a belief is justified involves determining whether certain beliefs are true. Internalism implies that justification can be recognized independently of considerations about truth. (For a somewhat different versions of the foregoing argument, and further discussion of how such arguments are related to internalism, see Goldman 1999 and Pappas 2014, section 8.)

Fortunately, however, the reliabilist has some room to maneuver. It isn't true that A has no way of establishing that Bi is true besides re-deploying the process P that produced Bi in the first place. Or at least, this isn't true in all cases. For one thing, if P is a perceptual process, it might be possible to obtain another token of Bi by deploying a different perceptual process. This would be the case if, say, the initial token of Bi results from a visual process that produces beliefs about shapes, and a second token results from a haptic process that is also concerned with shapes. But, also, additional tokens of Bi could be produced by processes that are not purely perceptual. Thus, a perceptual belief that a certain object is a torus can receive support from a belief produced by memory, and also from a belief produced by

attending to testimony. Again, a token of Bi might be produced by deducing it from other beliefs. And it might happen that the initial token of Bi contributes to the formation of a theory that turns out to generate accurate predictions. Appreciating that a token of Bi led to these predictions could lead to another token of Bi. Here the process would be the kind of reverse inference that is familiar from discussions of confirmation. Assuming that these additional processes are all reliable, they could provide A with justification for the belief-type **Bi** of which all of the additional tokens are instances, thereby providing support for **Bi** that is independent of the process P that produced the first instance. This shows that it is possible to use the belief-type **Bi** to assess the reliability of P even though P is responsible for one of its tokens.

This line of thought can be summarized by saying that it is possible to establish the reliability of a given belief-forming process by deploying other such processes in our reasoning, given that those others are themselves reliable. It is tempting to conclude that process reliabilism can accommodate the view that a theory of justification must allow for the fact that we can often recognize when beliefs are justified. Before we can draw this conclusion, however, we must take account of a concern. The reliability of a process P can be established by deploying other processes, but what happens when we need to determine that those other processes are reliable? It seems that we must deploy P at some point in our reasoning. After all, it is clear that we do not possess an infinite series of belief-forming processes. But if we must deploy P in establishing the reliability of other processes, won't our attempt to show that P is reliable be vitiated by circularity?

Let us take stock. We are considering whether it is possible to determine, without engaging in circular reasoning, which of one's belief-forming processes are reliable and therefore justification-conferring. This can be rephrased as the question of whether it is possible to use non-circular reasoning to form justified beliefs as to which of one's processes are reliable. The answer that is emerging is that there may well be a proper subset of reliable processes P1, . . . , Pn that can be used to establish the reliability of a broad range of other processes, but that P1, . . . , Pn cannot themselves be *shown* to be reliable without appealing to those other processes and therefore lapsing into circularity. The reliability of P1, . . . , Pn has to be *presupposed* or taken for granted; it can't be *established*. (If one tried to establish it, by appealing to a process Q that had been vetted by P1, . . . , Pn, one's effort would be defeated by the knowledge that in relying on Q one was engaging in circular reasoning. More specifically, if one tried to establish the reliability of P1, . . . , Pn by appeal to Q, one would be thwarted by a pair of beliefs—the belief that one had used P1, . . . , Pn in establishing the reliability of Q, and the belief that circular reasoning cannot provide justification.)

Where does this leave us? We have a tentative positive answer to our question, albeit a partial one. It may well be possible to form justified beliefs concerning the reliability of a large number of belief-forming processes. That is, arguably, we can

recognize reliability in a number of cases. As for the other processes, P1,…, Pn, their reliability has to be presupposed.

I now add that this is par for the course. Where T is any epistemological theory, there will generally be a distinction between the basic or foundational principles of T and the principles that are derived from the basic ones. The correctness of the basic principles will be presupposed. To be sure, the presuppositional status of the basic principles will sometimes be masked by claims to the effect that they are self evident, or by claims to the effect that they are constitutive of the concept of justification. But claims of these sorts can be challenged and almost always are. When challenged, they will have to be defended by arguments that depend on different principles. Sooner or later the defense will arrive at principles that are simply presupposed. "[J]ustifications must come to an end somewhere." (Nagel 1979, p. 12)

It is a familiar complaint about reliabilism that it fails to provide us with a set of "rules for the direction of the mind" (to borrow a phrase from Descartes 1955)—a set of instructions telling us how to manage the business of obtaining justified beliefs. It is maintained that only internalist theories can do justice to our need for such rules. But if "obtaining justified beliefs" means "obtaining beliefs that are epistemically justified," then the foregoing discussion shows that this criticism misses the mark. There may very well be procedures for establishing reliability. These procedures will be limited in scope, but other epistemological theories share this limitation.

VI. Reliabilism and the Epistemic Role of Perceptual Experience

This section is concerned with the justificatory roles that reliabilism assigns to conscious perceptual states—that is, to current perceptual experiences, to memories of very recent perceptual experiences, and to memories of less recent perceptual experiences that have recently been activated. There are perceptual states of many other types—for example, perceptual states that figure in early vision and perceptual memories that are in cold storage. But I wish to set such non-conscious states aside and focus on processes that take conscious perceptual states as inputs. We must focus on processes with conscious inputs and conscious outputs if we are to explain how agents are able to recognize processes as justification-conferring.

According to reliabilism, perceptual experiences contribute to epistemic justification by serving as the inputs to reliable belief-forming processes. This principle is prima facie plausible. Moreover, its plausibility is secured when we observe that it sustains our intuitions about the justifiedness of judgments that are triggered by experiences. That is, in cases in which we intuitively assess judgments as justified by experiences, there are in fact reliable processes leading from the experiences to the judgments. I will illustrate this point by considering cases of four kinds.

I will begin by asking how current perceptual experiences contribute to the justification of judgments that attribute appearance properties to objects. As we know from Chapter 2, the contents of perceptual experiences are appearance properties. The probability that an agent is actually confronted by an external object with appearance properties corresponding to the properties that the experience represents will inevitably be high. It will only be diminished by the small percentage of cases in which an agent is undergoing a hallucination and there is no corresponding object. Cases of illusion have no bearing on the probability, since an illusion consists of an external object and appearance properties that the object actually has. It follows that, as a matter of fact, almost all the inputs to the processes that generate beliefs about appearances will be veridical. Further, given this point, the reliability of a process P that takes current experiences as inputs and yields judgments about appearances as outputs must be extremely high. In view of this we can conclude that, according to reliabilism, judgments attributing appearances to objects have very high degrees of epistemic justification. This is of course fully in keeping with intuition.

What about perceptual judgments attributing objective sizes, objective shapes, objective forms of motion, and other low-level objective properties to external objects? How reliable are the processes running from perceptual experiences to judgments concerning properties of these kinds? It might seem that the relevant reliability ratings must be low because the appearance properties represented by experiences always underdetermine the objective factors that correspond to them. For example, since apparent sizes are determined by objective sizes acting in concert with objective distances, there is an infinite number of objective sizes that can give rise to any given apparent size. Despite this initial impression, however, a closer look reveals that the reliability ratings in question will actually be quite high. This is so for two reasons. First, perceptual experiences record more than the appearance properties of the objects that fall within the spotlight of attention. To illustrate, in addition to the apparent size of a birch tree on the horizon, the totality of the appearances registered by an experience of the tree will include a number of factors relevant to determining the objective distance of the tree, including aerial perspective, convergence of parallels, occlusion, position relative to the horizon, shading and shadows, and texture gradients. Moreover, if we consider not just the current experience of the tree but memories of the experiences immediately preceding it, the represented features will include appearances that reflect movement of the agent relative to the tree, thereby adding data concerning motion parallax to the available information about distance. There will also be distance-relevant appearance properties of the object itself, such as lack of visible fine-grained mereological structure. All in all, the contents of experiences that serve as inputs to the process that produces perceptual judgments about objective size will be quite rich. This brings me to the second reason why this process will have a high reliability rating. The process will involve many subprocesses that are capable of registering and assessing the relevance of all the

information concerning distance that the input experiences encode. Collectively, these subprocesses will place substantial constraints on unconscious estimates of objective distances, thereby ensuring that the upper levels of the encompassing process will deliver judgments about objective sizes that are approximately correct (though they will no doubt be pretty vague).

Some concepts are associated with perceptual templates or prototypes but are partially defined by theories. Biological kind concepts such as *chipmunk* and *armadillo* are paradigmatic members of this category. In normal cases we rely exclusively on perceptual information in deploying such concepts, but applications based on perception will be withdrawn if it is learned that the parents of a presented animal belonged to a different species (Keil 1989, Gelman 2003). Thus, application of concepts of biological kinds depends in part on abstract principles linking them to concepts having to do with parentage. What does process reliabilism tell us about the justification of judgments involving kind concepts of animals? It tells us that we must distinguish between two types of case. In cases of the first type, the relevant process takes perceptual experiences as inputs and yields judgments attributing kind membership as outputs. This process is quite reliable, though not perfectly so, assuming that the associated perceptual templates are sufficiently discriminating. This is because presented animals only rarely have rich sets of perceptible features that fail to reflect their parentage. In cases of the second type, the processes have two kinds of input: perceptual experiences of presented animals and experiences of other kinds, usually produced by testimony, that carry information concerning parentage. In the latter case, the degree of justification of the resulting judgment will depend in part on the truth or falsity of the testimony about parentage. When the testimony is true, the reliability rating of the process producing a kind-attribution will be high.

Let us finally consider judgments containing concepts that are embedded in empirical theories. Examples include judgments containing *electron* and judgements containing *gene*. Such judgments can be triggered by perceptual experiences—most notably by experiences of the outcomes of experiments. We can suppose that the belief-forming processes involved in such cases are in some sense shaped by the beliefs that are constitutive of the relevant theory. When this happens, the resulting judgments will probably be true if these theoretical beliefs are true. Hence, according to reliabilism, there will be many cases in which judgments involving very high level theoretical concepts are justified—a conclusion that coheres nicely with our intuitions about justification.

VII. Internalist Accounts of Justification and Rationality

I turn now to internalist theories of how to form and update beliefs. Theories that belong to this family have often been put forward as accounts of epistemic

justification (e.g., Chisholm 1977, BonJour 1985, Pollock 1986, Smithies 2019), but they are also sometimes seen as accounts of a different feature of beliefs that is known as *epistemic rationality* or more simply as *rationality* (Kelly 2003, Gupta 2019). I will adopt the labels "rationality" and "rational" here. Beliefs can be said to be rational, but so can whole systems of belief, agents who possess beliefs, and procedures or processes that produce beliefs.

In this section I will consider what two prominent internalist theories say about the role that experience plays in justification and rationality. One goal is simply to acknowledge that there are viable internalist theories that say interesting things about perceptual experience. I haven't meant to exclude internalist doctrines from the class of theories that make valuable claims about doxastic virtues. The second goal is to strengthen my case against internalist theories of epistemic justification by showing that, however successful the two theories may be as accounts of rationality, they fail as theories of epistemic justification.

The theories I will consider are *dogmatism* and *explanationism*.

As it was originally presented by Jim Pryor (Pryor 2000), dogmatism is concerned exclusively with the relationship between perceptual experience and judgments. In Pryor's account, it consists of the following principle, which I will call (PP):

(PP): Whenever you have an experience as of P's being the case, you thereby have immediate (*prima facie)* justification for believing that P. (Pryor 2000)

According to Pryor, a justification for believing P is immediate if it doesn't rest on any justification or evidence for believing other propositions. A perceptual justification for believing P is only prima facie because it is in principle possible for perceptual justifications to be *defeated* by prior beliefs of the relevant agent. Thus, prima facie perceptual justifications are subject to defeat by beliefs that contradict the beliefs that the justifications support (*rebutting defeaters*), and also by beliefs to the effect that the relevant belief-producing mechanisms are untrustworthy (*undercutting defeaters*). If a belief that is prima facie justified has no defeaters, the believer has an all things considered justification (an *ultima facie* justification) for believing P. (The theory of rebutting and undercutting defeaters is due to John Pollock. See, e.g., Pollock 1986.)

As Pryor points out, there are a number of different restrictions that might be placed on PP by placing constraints on the values of the variable P. To illustrate, it might be (i) required that P be concerned only with appearance properties, or (ii) required that P be concerned only with appearances and low-level objective properties, such as sizes, shapes, and colors, or (iii) required that P be concerned only with appearances and low-level objective properties and affordances, such as *climbable* and *graspable*.

Although Pryor uses the term "justification" in stating (PP), which suggests that he is thinking of it as a principle of epistemic justification, his proposal is

often understood as a principle of rationality. I will remain neutral between these interpretations. What is clear is that Pryor did not intend (PP) as a component of a reliabilist account of epistemic justification. It is supposed to be a foundational, free-standing principle, not a consequence of a more general reliabilist thesis. Moreover, it is meant to be an internalist principle. Hence, if we take it to be a proposal concerning epistemic justification, we must understand it to be a competitor of reliabilism, not a special case of it. In a few pages I will discuss the question of whether it is a successful competitor.

It is a principal goal of the proposal to block skeptical questions about transitions from perceptual experiences to beliefs. (PP) tells us that transitions of this sort cannot be legitimately questioned unless the skeptic can provide agents with genuine defeaters for them—that is, with reasons for believing that such transitions are untrustworthy. To do this, it isn't enough simply to allude to skeptical hypotheses. Rather, one must give reasons to think that skeptical hypotheses are true. Skeptics have no such reasons, nor do they claim that they do. They never claim that their hypotheses are true—just that they aren't refutable. (PP) implies that this isn't enough (Pryor 2000, 2013).

I turn now to explanationism. This is the view that it is rational to hold beliefs that provide good explanations for sets of established data, provided only (i) that there are no defeaters, and (ii) that there are no better explanations in the offing. Now in one sense of "explain," the propositions that comprise an explanation must be true; but explanationism presupposes a second sense in which explanatory relevance depends on features other than truth. More specifically, on this second sense, whether a set of propositions counts as an explanation for a set of data depends only on structural relations between the propositions and the data, and the goodness of an explanation depends on features of the propositions such as simplicity, predictive fecundity, and continuity with prior beliefs. The structural relations in question are principally deducibility and probabilification (in a subjective sense of "probability," according to which the probability of a proposition is a matter of an agent's degree of belief in the proposition). Hence, explanations in the intended sense of "explanation" depend only on factors that are internal to agents. Among other virtues, explanationism offers an account of how experiential data are relevant to the rationality of beliefs. The idea is that if a set of propositions provides a good explanation of a set of data, the data thereby make it more rational to believe the propositions. It is important to note that explanations of the relevant sort are not explanations of why the experiences in the data set are veridical, but rather of why the experiences have occurred. As an internalist theory, explanationism can have no involvement with truth or veridicality. To be sure, an explanation of the occurrence of an experience would consist of hypotheses about the external process that caused the experience, but what the explanation would establish is that it is rational to believe the hypothesis, not that the hypothesis is true. (Expositions and defenses of explanationism

can be found in Bonjour 1985, Lycan 1988, Vogel 1990, McCain 2014, Poston 2014, and McCain and Poston 2017.)

Let us now consider whether either of these internalist views can supplant reliabilism as an account of epistemic justification.

First, we should note that dogmatism has a component that corresponds to a component of reliabilism. To see this, consider the restricted version of (PP) that is concerned only with experiential judgments—that is, with judgments that are exclusively concerned with perceptible properties of external objects. The principle I have in mind looks like this:

(PPE): Whenever you have an experience as of an external object qua characterized by perceptible properties P1,…, Pn, you thereby have immediate prima facie justification for believing that the object is characterized by P1,…, Pn.

(PPE) corresponds to a belief-forming process that is highly reliable—namely, the process that takes experiences as inputs and yields perceptual judgments with the same contents as outputs. Thus, dogmatism vicariously specifies a reliable process without characterizing it in reliabilist terms. Accordingly, it can seem that dogmatism can explain many of the intuitions about justification that reliabilism explains without invoking truth or reliability. By the same token, it might seem that dogmatism is in a position to supplant reliabilism as an account of epistemic justification. As I see it, however, we should instead suppose that (PP) owes its plausibility to the fact that the inferential process it specifies is reliable. Where R is any reliable process, it will always be possible to describe R without mentioning the fact that it is reliable. In effect, it will always be possible to describe R in internalist terms. Because of this, it will always be possible to formulate internalist principles that seem to honor intuitions about epistemic justification as well as the corresponding reliabilist principles do. This is, I suggest, why Pryor's account of rational perceptual belief has sometimes been regarded as part of an internalist theory of *epistemic justification*. To appreciate the merits of this suggestion, suppose that a special case of the process described by (PPE) turns out to be unreliable—perhaps we get evidence from science which convinces us that color beliefs don't track objective properties and are therefore unreliable. News of this unreliability would call for restricting (PPE) in such a way as to exclude processes running from color experiences to color beliefs. This shows, I suggest, that our faith in (PPE) is conditional on a general presumption of reliability. We believe that the process it describes is generally reliable, and this explains our affection for it.

Another, closely related reason why dogmatism can't supplant reliabilism as a theory of epistemic justification is that we can't scale (PPE) up so as to enable dogmatism to account for a broader range of intuitions about doxastic merit. (PPE) is only concerned with purely perceptual beliefs. Reflection shows that there is no

simple, straightforward companion principle that can explain the justification of beliefs about higher-level properties. To see this, consider (PPH):

(PPH) Whenever you have an experience as of an external object qua characterized by high-level properties H1,…, Hn, you thereby have immediate prima facie justification for believing that the object is characterized by H1,…, Hn.

Now, when we speak of an experience as being *as of* a higher-level property—say, the property of belonging to a certain natural kind—what we must mean is that the experience has activated a concept of that property. Strictly speaking, the proprietary contents of experiences are restricted to *perceptible* properties. In view of this point, principle (PPH) is equivalent to the claim that if an experience causally triggers a concept of the property H, then you have immediate prima facie justification for believing the proposition that an object before you possesses H. And this claim is clearly quite wrong. All kinds of concepts can be triggered by perceptual experiences, including concepts that reflect bias and/or confusion and/or theories that are improperly grounded in experience. Thus, to scale (PPE) up to a believable form of (PPH), it would be necessary to add qualifications and restrictions to (PPH) that went well beyond those required for (PPE) itself. Now, if we were faced with such a principle, we would have to ask what motivated it. Given its Byzantine complexity, it would be very implausible that the principle was a ground-floor principle, motivated by a direct intuition. On the other hand, it would be quite plausible that the built-in restrictions were needed to block unreliable inferences, and that the positive portion of the principle was admissible because the inferences it licensed were reliable. In short, it would owe any standing it had as a guide to forming beliefs to considerations of reliability. This would be the best explanation of its appeal.

The central tenet of explanationism is the principle known as *Inference to the Best Explanation* (IBE):

(IBE) If (i) a set of propositions S stands in the appropriate explanatory relations to a set of data D (e.g. the members of D are deducible from S or are probabilified by it, where "probabilify" is understood in subjective terms), (ii) the members of S satisfy the appropriate constraints involving simplicity, predictive fecundity, and continuity with prior beliefs, and (iii) there is no better explanation of D, then it is prima facie rational to believe the members of S (on the basis of D).

As many authors have pointed out, (IBE) seems to play an important role in scientific reasoning, and its use seems to have led to a number of theories that can reasonably be regarded as true. If this is the case, then the principle corresponds to a belief-forming process that is at least moderately reliable—presumably reliable enough to warrant our trust. But (IBE) makes no mention of reliability! It is

an internalist principle. Moreover, the scope of this belief-forming process is extremely broad, with the result that the process can account for a large portion of the cases in which beliefs count intuitively as epistemically justified. Consider, for example, the proposition that the bark on the tree over there is dark brown. This proposition contributes to an explanation of the fact that the tree looks dark to me. Accordingly, the fact that the belief is justified is fully explained, even though reliability does not figure in the explanation. It might seem, therefore, that by promulgating (IBE), explanationists are in a position to account for a huge percentage of the intuitions about doxastic merit that reliabilism explains, and that explanationism is therefore capable of supplanting reliabilism, or at least, of assuming a large part of its role.

I have for some years been an admirer of explanationism (Hill 2013, chapter 1). Among other virtues, it offers a convincing internalist reply to skepticism. To amplify, explanationists have long maintained that commonsense explanations of our perceptual experiences are superior to the explanations offered by skeptical hypotheses (BonJour 1985, Vogel 1990, McCain and Poston 2017). This seems right. Indeed, so far at any rate, the explanatory value of skeptical hypotheses is within epsilon of zero. Descartes told us nothing about the inner workings of his Evil Genius, and Putnam had nothing to say about the computer program that controls the experiences of his brain in a vat. Now, of course, a skeptic might acknowledge this and say, "It doesn't matter that I can't actually state a skeptical hypothesis with the explanatory power of your commonsense beliefs. It's clearly possible, at least in the sense of 'logically possible,' for there to be skeptical hypotheses that fill the bill." But it does matter! We need to be able to predict and explain our perceptual experiences, and explanationism says that it is rational for us to accept the hypotheses that do the best job of filling this need. Hypotheses that are beyond our ken aren't even in the running. It doesn't matter that they exist somewhere in logical space.

To be sure, the programs for advanced virtual reality devices bring us closer to skeptical hypotheses that provide detailed explanations of experience. But by "closer" I mean "marginally closer." So far these programs have not done much to close the gap. (There is quite a difference between programs for cheesy video games and clunky museum tours, on the one hand, and programs that could explain the full panoply of perceptual experience, on the other.) Accordingly, today, and no doubt for some considerable time in the future, there is no real alternative to the explanation of experience that is offered by scientifically enlightened common sense. Note that it would be inadequate to invoke a program by saying something like "a VR program that is like what we have now but vastly more complicated."

Anyway, what is more important, there are supraempirical virtues that contribute to the goodness of explanations. One of these virtues is conservativeness—continuity with past and present theories. Another, related virtue is plausibility.

Skeptical hypotheses have no chance of possessing these virtues. But could their failure to be conservative and plausible somehow be balanced by their being simpler than our commonsense model of the world? It's hard to see how they could be simpler given that they must have at least as much explanatory power.

Offering an internalist answer to skepticism is not the only virtue of explanationism. Even so, however, it would be a mistake to suppose that it could become a serious competitor of reliabilism. Seen as a theory of epistemic justification, it can account for many of the intuitions that support reliabilism, but there are others that are forever beyond its reach. One problem is a special case of a problem we have already found to bedevil all internalist theories: explanationism cannot account for the importance that we attach to epistemic justification. We care much more about the property *produced by a reliable belief-forming process* than about the property *part of a best explanation of a set of data (in the non-factive sense of "explanation")*. If a belief possesses the first property, actions based on it ipso facto have some tendency to succeed. The latter property carries no such guarantee. There is also a second problem. If it is to be useful, epistemic justification must be a property to which we have ready access. Indeed, at least in the case of perceptual beliefs, we must be able to recognize it on the fly. (Is my hunting partner's belief that he just saw a seal justified?) But the property *part of a best explanation of a set of data* does not come close to meeting this condition. Constructing explanations of sets of data and showing that they have various virtues is a laborious process that requires more time and energy than are usually available.

I have claimed several times now that we care deeply about epistemic justification, so it is perhaps worth emphasizing that this really is so. Since it is agreed on all sides that justification is part of knowledge, we can measure its importance by considering the importance of knowledge. And we all know that knowledge is very important. When we are considering whether to act on a proposition P, we are likely to ask ourselves, "Do I really know whether P is true?" And when we seek advice from others about how to act or what to believe, we go looking for someone with expertise in the relevant area. Thus, when in need of directions to the freeway, we ask a passerby, "Do you know where the entrance to the freeway is?" And when we we're looking for quick help with a history essay, we might ask a roommate, "Do you know when the Oklahoma land rush occurred?" (Many philosophers hold that "knowledge is the norm of practical reasoning," meaning thereby that a proposition is eligible to serve as a premise in practical reasoning just in case the proposition is known to be true. Not every philosopher endorses this view. I for example do not. But those of us who reject it seem to be committed to an alternative thesis that makes the importance of epistemic justification even more vivid—namely, the thesis that epistemic justification is the norm of action (Hill and Schechter 2013).)

VIII. A Puzzle and Its Solution

In his rewarding book *Conscious Experience* (Gupta 2019), Anil Gupta maintains that questions about the role that experience plays in rationality can be pursued independently of metaphysical and scientific inquiries into the nature of experience. There is a case for this view. After all, norms are concerned with obligation and permission, and it is a familiar thought that how things are actually arranged may diverge, even radically, from how things ought to be or how they are permitted to be. But there is also a countervailing thought, which is that we must understand what experience *is* before we can determine what contributions to rationality it is capable of making. Understanding what experience is requires extensive metaphysical and empirical inquiries—the more the better.

So there is a question about whether, and if so to what extent, normative inquiries into rationality and metaphysical/scientific inquiries can be pursued independently. There is also a question of relative authority. Many naïve realists believe, or seem to believe, that our norms of rationality presuppose a substantive metaphysical picture of the nature of perceptual experience. They also seem to hold that we are required by the norms to accept the picture. Thus, we often find members of this school maintaining that our rational commitments presuppose that experience constitutively involves a relation of acquaintance, of the sort discussed in earlier parts of this book, and that these commitments are a sufficient reason for believing acquaintance actually exists. It is hard not to have some sympathy with the second part of this view. After all, if the norms of rationality presuppose a certain metaphysical picture, it can seem that we would be abandoning our claims to have rational beliefs if we were to reject that picture. But there is also another side to this story, which comes to the fore when we ask, "What happens if the metaphysical picture of experience that is presupposed by our norms of rationality is false?" The answer is that our sense that our beliefs are rational would be an illusion. This is given poignancy by the fact that our norms of rationality arguably presuppose the picture of experience that is contained in folk psychology. Folk psychology is like human nature: it is a mixture of admirable and deplorable components.

It turns out, then, that there are clashing perceptions of methodological independence and dependence, and also conflicting ideas concerning relative authority. Is there any way of getting beyond these paradoxical findings?

I believe so, though it involves making a strong assumption about the nature of rationality. The assumption is that the norms of rationality allow methods of inquiry that have proved empirically successful over a long run to count as rationality-conferring. Assuming that this is correct, the norms in question imply that it is advisable to rely on scientific method in empirical inquiries, and advisable to rely on the findings of science in evaluating metaphysical theories,

including the metaphysical picture of perceptual experience that is presupposed by our doxastic norms. On this view, our doxastic norms have a flexibility that permits revisions—not revisions of the fundamental norms themselves, but of the derivative norms that come from combining the fundamental norms with metaphysical pictures, such as the one that consists of the laws of folk psychology. That is, on this view, our doxastic norms have a fundamental component that, together with facts about the track records of methods of inquiry, permit scientific facts to call for revision of metaphysical pictures that are presupposed by the norms. On this picture, such presuppositions would always be provisional, hostage to discoveries of successful methods of inquiry that might eventually lead to facts that challenge them.

I won't attempt to give a full defense of this conception of our norms of rationality here. Instead I will just point out that it is plausible in view of the praise we bestow on scientific discoveries, even when they challenge long-held beliefs and long-respected inferences. Independently of that, I like the picture because, as noted, I believe that some of our current doxastic norms may presuppose highly dubious portions of folk psychology (one being the idea that there can be no appearance/reality distinction defined over perceptual qualia). The conception implies that it is possible to honor our basic doxastic norms while revising these folk presuppositions. Bravo!

I am concerned here with rationality rather than epistemic justification. Process reliabilism is clearly compatible with the idea that our derived epistemic norms are revisable in light of metaphysical and scientific discoveries, so the puzzle of methodology and authority doesn't arise for it. But the puzzle strongly suggests that we ought to adopt a more flexible attitude towards the norms of rationality than is our usual wont. I should add that in talking above about scientific facts that pose challenges to metaphysical pictures, I am referring to internalist facts involving the success and failure of predictions of experiences. (For more discussion of the idea that doxastic norms are revisable, see Field 2009.)

IX. Conclusion

If one wants to understand the epistemic role of perceptual experience, or that of perception generally, it is necessary to trace the boundaries of the epistemic domain. It is also necessary to survey the domain's internal geography. I have tried to achieve these ends by describing and defending process reliabilism. There are several other versions of externalism, but I have urged that reliabilism makes an especially good showing. I suggest that it is reasonable to commit to reliabilism, at least tentatively, and to the claims about the epistemic role of perception that it endorses.

Internalist theories may live on as accounts of a different doxastic virtue, rationality, but it would be a mistake to think of that virtue as bearing an important relation to knowledge. Hence, while it may be appropriate to say that internalist theories are concerned with a kind of rationality, we should not regard them as accounts of something that can legitimately be called *epistemic* rationality. (By the same token, it is a misnomer to speak of Bayesian *epistemology*.)

Are all internalist theories concerned with the same form of rationality? Are they in competition with each other? It seems that they must be concerned with different doxastic virtues, for when we consider the differences in structure among, say, explanationism, coherentism, Bayesianism, Gupta's theory (Gupta 2019, Hill forthcoming a), and Smithies theory (Smithies 2019), we find that they are quite radical. Their proponents pretty clearly have different ends in view. Are the forms of rationality with which these internalist theories are concerned important virtues, in the sense of conferring practical or prudential value on the beliefs that possess them? In some cases, at any rate, the answer is "yes." For example, it is valuable to possess a good explanation of a set of data, provided that the explanation is undefeated, even if the propositions that figure in the explanation are false. What matters is just that they are not known to be false. Having an undefeated explanation of data carries with it a sense of understanding, and a sense of understanding is psychologically rewarding. So the form of rationality that is captured by explanationism is worthwhile, even though the sense that one possesses a correct explanation is illusory much of the time. When we reflect, however, we find that the properties of conferring a sense of understanding pale in comparison with the property of being produced by a reliable process, and that the same is true of the beneficial properties of other internalist theories. (For example, the beneficial property celebrated by Bayesianism, *being in a position to avoid Dutch books*, is not nearly as valuable as the property *having beliefs that are produced by reliable processes*. To mention just one difference, the former property is negative, serving only to protect agents from a range of bad choices in a limited range of cases, while the latter property is positive, in that possessing it increases the odds of success in a broad range of ventures. For discussion see Vineberg 2016.)

References

Adair, E. R., Stevens, J. C., and Marks, L. E. (1967). "Thermally Induced Pain, the DOL Scale, and the Psychophysical Power Law," *American Journal of Psychology* 81: 147–64.

Åhs, F., Miller, S. S., Gordon, A. R., Lundstrom, J. N. (2013). "Aversive Learning Increases Sensory Detection Sensitivity," *Biological Psychology* 92: 135–41.

Altmann, C. F., Ono, K., Callan, A., Matsuhashi, M., Mima, T., and Fukuyama, H. (2013). "Environmental Reverberation Affects Processing of Sound Intensity in Right Temporal Cortex," *European Journal of Neuroscience* 38: 3210–20.

Ariew, A., Cummins, R., and Perlman, M., eds. (2002). *Functions*. Oxford: Oxford University Press.

Armstrong, D. M. (1980). *The Nature of Mind and Other Essays*. Ithaca, NY: Cornell University Press.

Arpaly, N., and Schroeder, T. (2014). *In Praise of Desire*. Oxford: Oxford University Press.

Aydede, M. (2019). "Pain," *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/pain/.

Baars, B. (1997). *In the Theater of Consciousness*. Oxford: Oxford University Press.

Badre, D. (2020). *On Task: How Our Brain Gets Things Done*. Princeton: Princeton University Press.

Baillargeon, R., Spelke, E., and Wasserman, S. (1985). "Object Permanence in 5-Month-Old Infants," *Cognition* 20: 191–208.

Bain, D. (2013). "What Makes Pains Unpleasant?" *Philosophical Studies* 166 (Supplement 1): s69–s89.

Bain, D. (2017). "Evaluativist Accounts of Pain's Unpleasantness," in J. Corns (ed.), *The Routledge Handbook of Philosophy of Pain*, pp. 40–50.

Bainbridge, K. E., Byrd-Clark, D., and Leopold, D. (2018). "Factors Associated with Phantom Odor Perception Among US Adults," *JAMA Otolaryngol Head Neck Surgery* 144(9): 807–14.

Baker, N., and Kellman, P. J. (2018). "Abstract Shape Representation in Human Visual Perception," *Journal of Experimental Psychology: General*, 147(9): 1295–308.

Barsalou, L. W. (1999). "Perceptual Symbol Systems," *Behavioral and Brain Sciences* 22, 577–609.

Basbaum, A. I. (2021). "Pain," in Kandel et al. (eds.), *Principles of Neural Science*, 6th edition, pp. 470–95.

Basbaum, A. I., and Jessell, T. M. (2013). "Pain," in Kandel et al. (eds.), *Principles of Neural Science*, 5th edition, pp. 530–55.

Batty, C. (2010). "What the Nose Doesn't Know: Non-Veridicality and Olfactory Experience," *Journal of Consciousness Studies* 17: 10–17.

Beck, J. (2018). "Analog Mental Representation," *WIREs Cognitive Science*, e1497, doi.org/10.1002/wcs.1479.

Beck, O. (2018). "Rethinking Naïve Realism," *Philosophical Studies* 176: 607–33.

Beck, O. (2021). "The Consciousness Knowledge Requires," paper delivered at a conference on New Waves in Relationalism, Madison, WI.

Berkley, G. (1953). *A New Theory of Vision and Other Writings*. New York: E. P. Dutton.

Bermudez, J. (2003). *Thinking without Words*. Oxford: Oxford University Press.
Bernstein, I. L. (1978). "Learned Taste Aversions in Children Receiving Chemotherapy," *Science* 200: 1302–3.
Berthier, M., Starkstein, S., and Leiguarda, R. (1988). "Pain Asymbolia: A Sensory-Limbic Disconnection Syndrome," *Annals of Neurology* 24: 41–9.
Biederman, I. (1987). "Recognition-by-components: A Theory of Human Image Understanding," *Psychological Review* 94: 115–47.
Biederman, I., and Ju, G. (1988). "Surface versus Edge-Based Determinants of Visual Recognition," *Cognitive Psychology* 20(1): 38–64.
Block, N. (1995). "On a Confusion about a Function of Consciousness," *The Behavioral and Brain Sciences* 18(2): 227–87.
Block, N. (1999). "Sexism, Racism, Ageism and the Nature of Consciousness," *Philosophical Topics* 26(1/2): 39–70.
Block, N. (2002). "Concepts of Consciousness," in D. J. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press, pp. 206–18.
Block, N. (2007). "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience," *Behavioral and Brain Sciences* 30: 481–99.
Block, N. (2011). "Perceptual Consciousness Overflows Cognitive Access," *Trends in Cognitive Science* 15: 567–75.
Block, N. (2015). "The Puzzle of Perceptual Precision," in T. Metzinger and J. M. Windt (eds.), *Open MIND*. Cambridge, MA: MIT Press, pp. 1–54.
Block, N. (2019). "What is Wrong with the No-report Paradigm and How to Fix It," *Trends in Cognitive Sciences* 23: 1003–13.
Block, N. (2022). *The Border between Seeing and Thinking*. Oxford: Oxford University Press.
BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA: Harvard University Press.
Borg, E., Harrison, R., Stazicker, J., and Salomons, T. (2020). "Is the Folk Concept of Pain Polyeidic?" *Mind & Language* 35: 29–47.
Bostrom, N. (2003a). "Are You Living in a Computer Simulation?" *Philosophical Quarterly* 53: 243–55.
Bostrom, N. (2003b). "The Simulation Argument: Why the Probability that You are Living in a Matrix is Quite High," *Times Higher Education Supplement*, June 16.
Bozzacchi, C., and Domini, F. (2015). "Lack of Depth Constancy for Grasping Movements in Both Virtual and Real Environments," *Journal of Neurophysiology* 114: 2242–8.
Bozzacchi, C., Volcic, R., and Domini, F. (2016). "Grasping in Absence of Feedback: Systematic Biases Endure Extensive Training," *Experimental Brain Research* 234: 255–65.
Brand, P., and Yancey, P. (1997). *The Gift of Pain*. Grand Rapids, MI: Zonderman.
Brascamp, J., Blake, R., and Knapen, T. (2015). "Negligible Fronto-parietal BOLD Activity Accompanying Unreportable Switches in Bistable Perception," *Nature Neuroscience* 18: 1672–8.
Braun, D. (1993). "Empty Names," *Noûs* 27: 449–69.
Brentano, F. (1973). *Psychology from an Empirical Standpoint*. New York: Humanities Press.
Broad, C. D. (1925). *The Mind and Its Place in Nature*, London: Kegan Paul.
Broggin, E., Savazzi, S., and Marzi, C. A. (2012). "Similar Effects of Visual Perception and Imagery on Simple Reaction Time," *The Quarterly Journal of Experimental Psychology*, 65(1): 151–64.
Bronkhorst, A. W., and Houtgast, T. (1999). "Auditory Distance Perception in Rooms," *Nature* 397: 517–18.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*, 2nd edition. Berkeley, CA: University of California Press.

Buckingham, G., Goodale, M. A., White, J. A., and Westwood, D. A. (2016). "Equal-magnitude Size-weight Illusions Experienced within and between Object Categories," *Journal of Vision* 16: 25.

Bülthoff, H. H., and Edelman, S. (1992). "Psychophysical Support for Two-dimensional View Interpolation Theory of Object Recognition," *Proceedings of the National Academy of Sciences USA* 89: 60–4.

Burge, T. (2010). *Origins of Objectivity*. Oxford: Oxford University Press.

Burge, J., and Burge, T. (forthcoming). "Shape, Perspective, and What Is and Is Not Perceived: Comment on Morales, Bax, and Firestone."

Bushnell, M. C., Villemure, C., and Duncan, G. H. (2004). "Psychophysical and Neurophysiological Studies of Pain Modulation by Attention," in Price and Bushnell (eds.), pp. 99–116.

Byrne, A. (2001). "Intentionalism Defended," *The Philosophical Review*, 110: 199–240.

Byrne, A. (2016). "Hill on Mind," *Philosophical Studies* 173: 831–9.

Byrne, A. (2018). *Transparency and Self-knowledge*. Oxford: Oxford University Press.

Byrne, A., and Logue, H. (2008). *Disjunctivism: Contemporary Readings*. Cambridge, MA: MIT Press.

Camp, E. (2009). "A Language of Baboon Thought?" in R. Lurz (ed.), *Philosophy of Animal Minds*. Cambridge: Cambridge University Press, pp. 108–27.

Campagnoli, C., Croom, S., and Domini, F. (2017). "Stereovision for Action Reflects our Perceptual Experience of Distance and Depth," *Journal of Vision* 17(9): 1–26.

Campbell, J. (2009). "Consciousness and Reference," in A. Beckermann, B. McLaughlin, and S. Walter (eds.), *The Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press, pp. 648–62.

Carey, S. (2009). *The Origin of Concepts*. Oxford: Oxford University Press.

Carey, S., and Xu, F. (2001). "Infants' Knowledge of Objects: Beyond Object Files and Object Tracking," *Cognition* 80: 179–213.

Carrasco, M. (2011). "Attention: The Last Twenty-Five Years," *Vision Research* 51: 1484–525.

Carrasco, M., Ling, S., and Read, S. (2004). "Attention Alters Appearance." *Nature Neuroscience* 7: 308–13.

Carrasco, M., Penpeci-Talgar, C., and Eckstein, M. (2000). "Spatial Covert Attention Increases Contrast Sensitivity across the CFS: Support for Signal Enhancement," *Vision Research* 40: 1203–15.

Carruthers, P. (2011). *The Opacity of Mind*. Oxford: Oxford University Press.

Carruthers, P. (2018). "The Causes and Contents of Inner Speech," in P. Langland-Hassan and A. Vicente (eds.), *Inner Speech: New Voices*. Oxford: Oxford University Press, pp. 31–52.

Casati, R., and Dokic, J. (1994). *La Philosophie du Son*. Nîmes: Chambon.

Casati, R., Dokic, J., and Di Bona, E. (2020). "Sounds," *Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/entries/sounds/.

Chalmers, D. J. (1997). *The Conscious Mind*. Oxford: Oxford University Press.

Chalmers, D. J. (2022). *Reality+*. New York: W. W. Norton.

Chang, L., and Tsao, D. (2017). "The Code for Facial Identity in the Primate Brain," *Cell* 169: 1013–28.

Cheney, D. L., and Seyfarth, R. M. (1990). *How Monkeys See the World*. Chicago: University of Chicago Press.

Cheyney, D. L., and Seyfarth, R. M. (2008). *Baboon Metaphysics*. Chicago: University of Chicago Press.

Chisholm, R. M. (1957). *Perceiving*. Ithaca, NY: Cornell University Press.
Chisholm, R. M., (1977). *Theory of Knowledge*, 2nd edition. Englewood Cliffs, NJ: Prentice Hall.
Choi, S., and Fara, M. (2018). "Dispositions," *Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/entries/dispositions.
Churchland, P. (1995). *The Engine of Reason, The Seat of the Soul*. Cambridge, MA: MIT Press.
Cohen, J., Hansel, C. E. M., and Sylvester, J. D. (1953). "A New Phenomenon in Time Judgment," *Nature* 172: 901.
Cohen, M. A., and Dennett, D. C. (2011). "Consciousness Cannot Be Separated from Function," *Trends in Cognitive Sciences* 15: 358–64.
Cohen, M. A., Dennett, D. C., and Kanwisher, N. (2016). "What Is the Bandwidth of Perceptual Experience?" *Trends in Cognitive Sciences* 20: 324–35.
Cohen, S. (1984). "Justification and Truth", *Philosophical Studies* 46: 279–96.
Coleman, P. D. (1962). "Failure to Localize the Source Distance of an Unfamiliar Sound," *Journal of the Acoustical Society of America* 34(3): 345–6.
Connolly, K. (2019). *Perceptual Learning*. Oxford: Oxford University Press.
Coren, S., and Ward, L. M. (1989). *Sensation and Perception*. New York: Harcourt, Brace, and Jovanovich.
Corley, M., Brocklehurst, P., and Moat, H. S. (2011). "Error Biases in Inner and Overt Speech: Evidence from Tongue Twisters," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37: 162–75.
Corns, J., ed. (2017). *The Routledge Handbook of the Philosophy of Pain*. New York: Routledge.
Craig, A. D. (2010). "The Sentient Self," *Brain Structure and Function* 214: 563–77.
Craig, A. D. (2015). *How Do You Feel?* Princeton: Princeton University Press.
Craig, A. D., Krout, K., and Andrew, D. (2001). "Quantitative Response Characteristics of Thermoreceptive and Nociceptive Lamina I Spinothalamic Neurons in the Cat," *Journal of Neurophysiology* 86: 1459–80.
Cummins, R. (1975). "Functional Analysis," *The Journal of Philosophy* 72: 741–65.
Cummins, R. (1989). *Meaning and Mental Representation*. Cambridge, MA: MIT Press.
Cummins, R. (1996). *Representations, Targets, and Attitudes*. Cambridge, MA: MIT Press.
Cutter, B., and Tye, M. (2011). "Tracking Representationalism and the Painfulness of Pain," *Philosophical Issues* 21: 90–109.
Damasio, A. R. (1999). *The Feeling of What Happens*. New York: Harcourt, Brace, and Compa.
Damasio, A. R. (2003). *Looking for Spinoza*. Orlando, FL: Harcourt, Inc.
Damasio, A. R. (2021). *Feeling and Knowing: Making Minds Conscious*. New York: Pantheon Books.
Daoust, L. (2021). "Stability by Degrees: Conceptions of Constancy from the History of Perceptual Psychology," *History and Philosophy of the Life Sciences* 43(1): 1–22.
Davidson, D. (1986). "A Coherence Theory of Truth and Knowledge," in E. LePore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*. Oxford: Basil Blackwell, pp. 307–19.
Davidson, D. (1987). "Knowing One's Own Mind," *Proceedings and Addresses of the American Philosophical Association* 60: 441–58.
Davies, M. (1992). "Perceptual Content and Local Supervenience," *Proceedings of the Aristotelian Society* 92: 21–45.
Dehaene, S. (2014). *Consciousness and the Brain*. New York: Penguin Books.
Dehaene, S., and Changeux, J. P. (2011). "Experimental and Theoretical Approaches to Conscious Processing," *Neuron* 70: 200–27.

Dennett, D. C. (1978). "Why You Can't Make a Computer That Feels Pain," *Synthese* 38: 415–56.
Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.
Dennett, D. C. (1988). "Quining Qualia," in A. J. Marcel and E. Bisiach (eds.), *Consciousness in Contemporary Science*. Oxford: Oxford University Press, pp. 42–77.
Dennett, D. C. (2005). *Sweet Dreams*. Cambridge, MA: MIT Press.
Descartes, R. (1955). *Rules for the Direction of the Mind*, in E. Haldane and C. R. T. Ross (eds.), *Philosophical Works of Descartes, Volume I* (New York: Dover, 1955).
Descartes, R. (1971). *Dioptrics*, in Elizabeth Anscombe and Peter T. Geach (eds.), *Descartes: Philosophical Writings*. Indianapolis, IN: Bobbs-Merrill Company.
Descartes, R. (2017). *Meditations on First Philosophy*, 2nd edition. Cambridge: Cambridge University Press.
Dewaele, J. M. (2011). "Self-reported Use and Perception of the L1and L2 among Maximally Proficient Bi- and Multilinguals: A Quantitative and Qualitative Investigation," *International Journal of the Sociology of Language*, 208: 25–51.
Dosher, B., and Lu, Z. (2020). *Perceptual Learning*. Cambridge, MA: MIT Press.
Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.
Dretske, F. (1986). "Misrepresentation," in R. Bogdan (ed.), *Belief: Form, Content, and Function*. Oxford: Oxford University Press, pp. 17–36.
Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press.
Dretske, F. (1995). *Naturalizing the Mind*. Cambridge, MA: MIT Press.
Driver, J., and Vuilleumier, P. (2001). "Perceptual Awareness and its Loss in Unilateral Neglect and Extinction," *Cognition* 79(1–2): 39–88.
Ducasse, C. J. (1952). "Moore's Refutation of Idealism," in P. A. Schilpp (ed.), *The Philosophy of G. E. Moore*. New York: Tudor.
Durgin, F. H. (2009). "When Walking Makes Perception Better," *Current Directions in Psychological Science* 18(1): 43–7.
Durgin, F. H. (2014). "Angular Scale Expansion Theory and the Misperception of Egocentric Distance in Locomotor Space," *Psychology and Neuroscience* 7: 253–60.
Durgin, F. H., and Li, Z. (2011). "Perceptual Scale Expansion: An Efficient Angular Coding Strategy for Locomotor Space," *Attention, Perception & Psychophysics* 73: 1856–70.
Durgin, F. H., and Li, Z. (2017). "Why Do Hills Look so Steep?" in Arthur G. Shapiro and Dejan Todorvic (eds.) *Oxford Compendium of Visual Illusions*. Oxford: Oxford University Press, pp. 190–7.
Engel, A. K., Fries, P., and Singer, W. (2001). "Dynamic Predictions: Oscillations and Synchrony in Top-down Processing," *Nature Reviews Neuroscience* 2(10): 704–16.
Epstein, W., and Park, J. N. (1963). "Shape Constancy: Functional Relationships and Theoretical Formulations," *Psychological Bulletin* 60(3): 265–88.
Ester, E. F., Sprague, T. C., and Serences, J. T. (2020). "Category Learning Biases Sensory Representations in Occipitoparietal Cortex," *The Journal of Neuroscience* 40(4): 917–31.
Evans, G. (1982). *Varieties of Reference*. Oxford: Oxford University Press.
Farennikova, A. (2013). "Seeing Absence," *Philosophical Studies* 166(3): 429–54.
ffytche, D. H. (2013). "The Hallucinating Brain: Neurobiological Insights into the Nature of Hallucinations," in F. Macpherson and D. Platchias (eds.), *Hallucination*. Cambridge, MA: The MIT Press, pp. 45–63.
Field, H. (2009). "Epistemology without Metaphysics," *Philosophical Studies* 143: 249–90.
Finke, R. A. (1989). *Principles of Mental Imagery*. Cambridge, MA: MIT Press.

Finke, R. A., and Pinker, S. (1983). "Directional Scanning of Remembered Visual Patterns," *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9: 398–410.

Firestone, C., and Scholl, B. (2016). "Cognition Does Not Affect Perception: Evaluating the Evidence for 'Top-down' Effects," *Behavioral and Brain Sciences* 39: 1–72.

Fleming, S. M. (2021). *Know Thyself: The Science of Self Awareness*. New York: Basic Books.

Flombaum, J. I., Kundey, S. M., Santos, L. R., and Scholl, B. J. (2004). "Dynamic Object Individuation in Rhesus Macaques: A Study of the Tunnel Effect," *Psychological Science* 15: 795–800.

Fodor, J. (2007). "The Revenge of the Given," in B. McLaughlin and J. Cohen (eds.), *Contemporary Debates in Philosophy of Mind*. Oxford: Blackwell, pp. 105–17.

Fodor, J. A. (1990). *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

Fodor, J. A. (2008). *LOT 2: The Language of Thought Revisited*. Oxford: Oxford University Press.

Foley, J. M. (1980). "Binocular Distance Perception," *Psychological Review* 87 (5): 411–34.

Foley, J. M., Ribeiro-Filho, N. P., and DaSilva, J. A. (2004). "Visual Perception of Extent and the Geometry of Visual Space," *Vision Research* 44: 147–56.

Foltz, E. L., and White, L. E. (1962). "Pain 'Relief' by Frontal Cingulotomy," *Journal of Neurosurgery* 19: 89–100.

Frank, M. C., Fedorenko, E., Lai, P., Saxe, R., and Gibson, E. (2012). "Verbal Interference Suppresses Exact Numerical Representation," *Cognitive Psychology* 64(1–2): 74–92.

Franklin, A. (2016). "Infant Color Categories," in L. Ronier (ed.), *Encyclopedia of Color Science and Technology*. Heidelberg: Springer-Verlag, pp. 765–70.

Franklin, A., Skelton, A., and Catchpole, G. (2014). "The Case for Infant Colour Categories," in W. Anderson, C. P. Biggam, C. Hough, and K. Christian (eds.), *Colour Studies: A Broad Spectrum*. London: John Benjamins Publishing Company, pp. 169–80.

Freedman, D. J., Riesenhuber, M., Poggio, T., and Miller, E. K. (2001). "Categorical Representation of Visual Stimuli in the Primate Prefrontal Cortex," *Science* 291(2): 312–16.

Freedman, D. J., Riesenhuber, M., Poggio, T., and Miller, E. K. (2003). "A Comparison of Primate Prefrontal and Inferior Temporal Cortices during Visual Categorization," *The Journal of Neuroscience* 23(12): 5235–46.

Frege, G. (1952). "On Sense and Reference," *Philosophical Review* 57(3): 209–30.

French, C., and Phillips, I. (2020). "Austerity and Illusion," *Philosopher's Imprint* 20(15): 1–19.

Freud, S. (1965). *A General Introduction to Psychoanalysis*. New York: Washington Square Press.

Frisby, J. P., and Stone, J. V. (2010). *Seeing*, 2nd edition. Cambridge, MA: MIT Press.

Fulkerson, M. (2014). *The First Sense: A Philosophical Study of Human Touch*. Cambridge, MA: MIT Press.

Garcia, J., and Koelling, R. A. (1966). "Relation of Cue to Consequence in Avoidance Learning," *Psychonomic Science* 4:123–4.

Garcia, J., Kimeldorf, D. J., and Koelling, R. A. (1955). "Conditioned Aversion to Saccharin Resulting from Exposure to Gamma Radiation," *Science* 122: 157–8.

Gardner, E. P. (2021). "Touch," in E. R. Kandel, J. D. Koester, S. H. Mack, and S. A. Siegelbaum (eds.), *Principles of Neural Science*, 6th edition. New York: McGraw-Hill, pp. 435–59.

Gardner, E. P., and Johnson, K. O. (2013). "Touch," in E. R. Kandel, J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum, and A. J. Hudspeth (eds.), *Principles of Neural Science*, 5th edition. New York: McGraw-Hill, pp. 498–529.

Gelman, S. (2003). *The Essential Child*. Oxford: Oxford University Press.

Gennaro, R. J. (2012). *The Consciousness Paradox*. Cambridge, MA: MIT Press.

Gentaz, E., and Hatwell, Y. (2004). "Geometrical Haptic Illusions: The Role of Exploration in the Müller-Lyer, Vertical–horizontal, and Delboeuf Illusions," *Psychonomic Bulletin and Review* 11(1): 31–40.
Gertler, B. (2011). *Self-Knowledge*. New York: Routledge.
Gibbard, A. (1975). "Contingent Identity," *Journal of Philosophical Logic* 4(2): 187–221.
Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*, Boston, MA: Houghton Mifflin.
Gilchrist, A. (1977). "Perceived Lightness Depends on Spatial Arrangement," *Science* 195: 185–7.
Gilchrist, A. (2006). *Seeing Black and White*. Oxford: Oxford University Press.
Godfrey-Smith, P. (1993). "Functions: Consensus without Unity," *Pacific Philosophical Quarterly* 74(3): 196–208.
Godfrey-Smith, P. (2020). *Metazoa*. New York: Farrar, Straus, and Giroux.
Goldfarb, L., and Treisman, A. (2013). "Counting Multidimensional Objects: Implications for the Neural-Synchrony Theory," *Psychological Science* 24(3): 266–71.
Goldman, A. I. (1979). "What Is Justified Belief?" in G. S. Pappas (ed.), *Justification and Knowledge: New Studies in Epistemology*. Dordrecht: Reidel, pp. 1–25.
Goldman, A. I. (1999). "Internalism Exposed," *The Journal of Philosophy* 96(6): 271–93.
Goldman, A. I., and Beddor, B. (2015). "Reliabilist Epistemology," *Stanford Encyclopedia of Philosophy* https://plato.stanford.edu/entries/reliabilism/#RepRefMod.
Goldman, A. I., and McGrath, M. (2015). *Epistemology*. Oxford: Oxford University Press.
Goldstein, E. B. (2010). *Sensation and Perception*, 8th edition. Belmont, CA: Wadsworth.
Goldstone, R. L., and Byrge, L. A. (2015). "Perceptual Learning," in M. Matthan (ed.), *The Oxford Handbook of the Philosophy of Perception*. Oxford: Oxford University Press, pp. 812–32.
Grabenhorst, F., and Rolls, E. T. (2011). "Value, Pleasure and Choice in the Ventral Prefrontal Cortex," *Trends in Cognitive Sciences* 15(2): 56–67.
Grahek, N. (2007). *Feeling Pain and Being in Pain*. Cambridge, MA: MIT Press.
Granrud, C. E. (2009). "Development of Size Constancy in Children: A Test of the Metacognitive Theory," *Attention, Perception, and Psychophysics* 71: 644–54.
Granrud, C. E. (2012). "Judging the Size of a Distant Object: Strategy Use by Children and Adults," in Gary Hatfield and Susan Allred (eds.) *Visual Experience: Sensation, Cognition, and Constancy*. Oxford: Oxford University Press, pp. 13–34.
Grant, J. A. (2014). "Meditative Analgesia: The Current State of the Field," *Annals of the New York Academy of Sciences* 1307 (1): 55–63.
Green, E. J. (2017a). "Psychosemantics and the Rich/Thin Debate," *Philosophical Perspectives* 31: 153–86.
Green, E. J. (2017b). "Attentive Visual Reference," *Mind & Language* 32(1): 3–38.
Green, E. J. (2018). "What Do Object Files Pick Out?" *Philosophy of Science*, 85: 177–200.
Green, E. J., and Quilty-Dunn, J. (2017). "What is an Object File?" *The British Journal for Philosophy of Science*. doi:10.1093/bjps/axx055
Guerrero, M. C. M. (2005). *Inner Speech—L2: Thinking Words in a Second Language*. New York: Springer.
Gupta, A. (2019). *Consciousness and Experience*. Cambridge, MA: Harvard University Press.
Haberman, J., Brady, T. F., and Alvarez, G. A. (2015). "Individual Differences in Ensemble Perception Reveal Multiple, Independent Levels of Ensemble Representation," *Journal of Experimental Psychology: General* 144(2): 432–46.
Hardcastle, V. (1999). *The Myth of Pain*. Cambridge, MA: MIT Press.

Harlow, P. J., Kim, J., and Anderson, B. L. (2102). "The Perception and Misperception of Spectral Surface Reflectance," *Current Biology* 22(20): 1909–13.

Harman, G. (1990). "The Intrinsic Quality of Experience," *Philosophical Perspectives* 4: 31–52.

Hartcher-O'Brien, J., Terekhov, A., Auvray M., and Hayward, V. (2014). "Haptic Shape Constancy across Distance," *Proceedings of the International Conference on Human Haptic Sensing and Touch Enabled Computer Applications.* Berlin: Springer, pp. 77–84.

Heathwood, C. (2018). "Unconscious Pleasures and Attitudinal Theories of Pleasure," *Utilitas* 30(2): 219–27.

Helson, H (1930). "The Tau Effect—An Example of Psychological Relativity," *Science* 71: 536–7.

Hermer-Vazquez, L., Spelke, E. S., and Katsnelson, A. S. (1999). "Sources of Flexibility in Human Cognition: Dual-task Studies of Space and Language," *Cognitive Psychology* 39(1): 3–36.

Hill, C. S. (1996). "Process Reliabilism and Cartesian Scepticism," *Philosophy and Phenomenological Research*, LV1: 567–81.

Hill, C. S. (2002). *Thought and World*. Cambridge: Cambridge University Press.

Hill, C. S. (2009). *Consciousness*. Cambridge: Cambridge University Press.

Hill, C. S. (2013). *Meaning, Mind and Knowledge*. Oxford: Oxford University Press

Hill, C. S. (2014a). "OW! The Paradox of Pain," in C. S. Hill, *Meaning, Mind, and Knowledge*. Oxford: Oxford University Press, pp. 155–76.

Hill, C. S. (2014b). "Locating Qualia: Do They Reside in the Mind or in the Body and the World?" in C. S. Hill, *Meaning, Mind, and Knowledge*. Oxford: Oxford University Press, pp. 177–96.

Hill, C. S. (2016). "Perceptual Relativity," *Philosophical Topics*, 44: 179–200.

Hill, C. S. (2017). "Fault Lines in Familiar Concepts of Pain," in J. Corns (ed.), *The Routledge Handbook of Philosophy of Pain*, pp. 60–9.

Hill, C. S. (2019). "Perceptual Existentialism Sustained," *Erkenntnis* doi.org/10.1007/s10670-019-00160-z.

Hill, C. S. (2020). "Appearance and Reality," *Philosophical Issues* 30 (2020): 175–91.

Hill, C. S. (2021). "What Is a Concept?" in R. G. Oliviera and K. J. Corcoran (eds.), *Common Sense Metaphysics*. New York: Routledge.

Hill, C. S. (forthcoming a). "Gupta Has Built a Magnificent Mansion, But Can We Live in It?" *Philosophy and Phenomenological Research*.

Hill, C. S. (forthcoming b). "Neander on a Mark of the Mental," *Philosophy and Phenomenological Research*.

Hill, C. S., and Pavese, C. (2021). "A Tribute to Karen Neander," *Biological Theory*, https://doi.org/10.1007/s13752-021-00383-w.

Hill, C. S., and Schechter, J. (2013). "Hawthorne's Lottery Puzzle and the Nature of Belief," in Hill 2013, pp. 253–72.

Hirsch, E. (1976). "Physical Identity," *Philosophical Review* 85: 357–89.

Huemer, M. (2001). *Skepticism and the Veil of Perception*. Lanham, MD: Rowman and Littlefield.

Hume, D. (2007). *An Enquiry Concerning Human Understanding*. Cambridge: Cambridge University Press.

Hume, D. (2011). *A Treatise of Human Nature*, in D. F. Norton and M. J. Norton (eds), *David Hume*, Volume I. Oxford: Clarendon Press.

Hurlbert A. (2003). "Colour Vision: Primary Visual Cortex Shows Its Influence," *Current Biology* 13: R270–2.

Hurlburt, R. T., and Schwitzgebel, E. (2007). *Describing Inner Experience? Proponent Meets Skeptic*. Cambridge, MA: MIT Press.

Hurlburt, R. T., Alderson-Day, B., Kühn, S., and Fernyhough, C. (2016). "Exploring the Ecological Validity of Thinking on Demand: Neural Correlates of Elicited vs. Spontaneously Occurring Inner Speech," PLOS ONE 11(2) DOI:10.1371/journal.pone.0147932.

Hyman, S. E., and Cohen, J. D. (2013). "Disorders of Mood and Anxiety," in Kandel et al. *Principles of Neural Science*, pp. 1402–24.

Ibañez, A., Gleichgerrcht, E., and Manes, F. (2010). "Clinical Effects of Insular Damage in Humans," *Brain Structures and Function* 214: 397–410.

Jackson, F. (1977). *Perception: A Representative Theory*. Cambridge: Cambridge University Press.

Jackson, F. (1986). "What Mary Didn't Know," *Journal of Philosophy* 83 (5): 291–5.

Jacobson, H. (2019). "Not Only a Messenger: Towards an Attitudinal-Representational Theory of Pain," *Philosophy and Phenomenological Research* 99(2): 382–408.

James, W. (1884). "What is an Emotion?" *Mind* 9: 188–205.

James, W. (1960). *The Will to Believe, Human Immortality, and Other Essays in Popular Philosophy*. New York: Dover Books.

Johnston, E. B. (1990). "Systematic Distortions of Shape from Stereopsis," *Vision Research* 31(7/8): 1351–60.

Joyce, J. (2003). "Bayes' Theorem," https://plato.stanford.edu/entries/bayes-theorem/

Kahneman, D., and Treisman, A. (1984). "Changing Views of Attention and Automaticity," in R. Parasuraman and D. R. Davies (eds.), *Varieties of Attention*. New York: Academic Press, pp. 29–61.

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). "The Reviewing of Object Files: Object-specific Integration of Information," *Cognitive Psychology* 24: 175–219.

Kandel, E. R., Koester, J. D., Mack, S. H., and Siegelbaum, S. A., eds. (2021). *Principles of Neural Science*, 6th edition. New York: McGraw-Hill.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., and Hudspeth, A. J., eds. (2013). *Principles of Neural Science*, 5th edition. New York: McGraw-Hill.

Kant, I. (2003). *Critique of Pure Reason*. New York: Palgrave Macmillan.

Kaplan, D. (1989). "Demonstratives," in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*. Oxford: Oxford University Press, pp. 481–565.

Kappers, A. M. L. (2007). "Haptic Space Processing—Allocentric and Egocentric Reference Frames," *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 61(3): 208–18.

Kappers, A. M. L., and Koenderink, J. J. (1999). Haptic perception of spatial relations. *Perception*, 28(6), 781–95.

Kass, M. D., Guang, S. A., Moberly, A. H., and McGann, J. P. (2016). "Changes in Olfactory Sensory Neuron Physiology and Olfactory Perceptual Learning After Odorant Exposure in Adult Mice," *Chemical Senses* 41: 123–33.

Kass, M. D., Rosenthal, M. C., Pottackal, J., and McGann, J. P. (2013). "Fear Learning Enhances Neural Responses to Threat-Predictive Sensory Stimuli," *Science* 342(6164): 1389–92.

Keil, F. C. (1989). *Concepts, Kinds, and Cognitive Development*. Cambridge, MA: MIT Press.

Kelly, T. (2003). "Epistemic Rationality as Instrumental Rationality: A Critique," *Philosophy and Phenomenological Research* 66(3): 612–40.

Kiani, R., Esteky, H., Mirpour, K., and Tanaka, K. (2007). "Object Category Structure in Response Patterns of Neuronal Population in Monkey Inferior Temporal Cortex," *Journal of Neurophysiology* 97: 4296–309.

Kim, J. (2006). *Philosophy of Mind*, 2nd edition. Boulder, CO: Westview Press.
Klatzky, R. L., and Lederman, S. J. (2003). "Representing Spatial Location and Layout from Sparse Kinesthetic Contacts," *Journal of Experimental Psychology: Human Perception and Performance* 29(2): 310–25.
Klein, C. (2015). "What Pain Asymbolia Really Shows," *Mind* 124: 493–516.
Klein, C. (2017). "Imperativism," in J. Corns (ed.), *The Routledge Handbook of Philosophy of Pain*, pp. 40–50.
Knight, D. C., Smith, C. N, Stein, E. A., and Helmstetter, F. J. (1999). Functional MRI of Human Pavlovian Fear Conditioning: Patterns of Activation as a Function of Learning," *Neuroreport* 10: 3665–70.
Koenderink, J. J., and van Doorn, A. J. (1977). "How an Ambulant Observer Can Construct a Model of the Environment from the Geometric Structure of the Visual Inflow," in G. Hauske and E. Butenandt (eds.), *Kybernetik*. Munich: Oldenbourg, pp. 224–47.
Kolarik, A. J., Moore, B. C. J., Zahorik, P., Cirstea, S., and Pardhan, S. (2016). "Auditory Distance Perception in Humans: A Review of Cues, Development, Neuronal Bases, and Effects of Sensory Loss," *Attention, Perception, & Psychophysics* 78(2): 373–95.
Kosslyn, S. (1980). *Image and Mind*. Cambridge, MA: Harvard University Press.
Kosslyn, S., Thompson, W. L., and Ganis, G. (2006). *The Case for Mental Imagery*. Oxford: Oxford University Press.
Kosslyn, S. M. (1973). "Scanning Visual Images: Some Structural Implications," *Perception and Psychophysics* 14: 90–4.
Kozuch, B. (2020). "No Pain, No Gain (in Darwinian Fitness): A Representational Account of Affective Experience," *Erkenntnis* 85: 693–714.
Kringelbach, M. L. (2005). "The Human Orbitofrontal Cortex: Linking Reward to Hedonic Experience," *Nature Reviews Neuroscience* 6: 691–702.
Kringelbach, M. L., and Berridge, K. C., eds. (2010). *Pleasures of the Brain*. Oxford: Oxford University Press.
Kringelbach, M. L., and Berridge, K. C. (2015). "Pleasures of the Brain," *Neuron* 86: 646–64.
Kringelbach, M. L., and Berridge, K. C. (2017). "The Affective Core of Emotion: Linking Pleasure, Subjective Well-Being, and Optimal Metastability in the Brain," *Emotion Review* 9(3): 191–9.
Kripke, S. (1981). Naming *and Necessity*. Oxford: Blackwell.
Kühn, S., Fernyhough, C., Alderson-Day, B., and Hurlburt, R. T. (2014). "Inner Experience in the Scanner: Can High Fidelity Apprehensions of Inner Experience be Integrated with fMRI?," *Frontiers in Psychology* 5 # 1393 doi: 10.3389/fpsyg.2014.01393.
Kulvicki, J. (2008). "The Nature of Noise", *Philosopher's Imprint*, 8(011): 1–16.
Lande, K. J. (2018). "The Perspectival Character of Perception," *The Journal of Philosophy* 115(4): 187–214.
Landman, R., Spekreijse, H., and Lamme, V. A. F. (2003). "Large Capacity Storage of Integrated Objects before Change Blindness," *Vision Research* 43: 149–64.
Langland-Hassan, P. (2018). "From Introspection to Essence: The Auditory Nature of Inner Speech," in Langland-Hassan and Vicente (eds.), *Inner Speech: New Voices*, pp. 78–104.
Langland-Hassan, P. (2020). "Inner Speech," *WIREs Cognitive Science*, #1544, doi:https://doi.org/10.1002/wcs.1544
Langland-Hassan, P., and Vicente, A., eds. (2018). *Inner Speech: New Voices*. Oxford: Oxford University Press.
Lau, H., and Rosenthal, D. (2011). "Empirical Support for Higher Order Theories of Conscious Awareness," *Trends in Cognitive Sciences* 15(8): 365–73.

Lederman, S. J., and Klatzky, R. L. (2009). "Haptic Perception: A Tutorial," *Attention, Perception, and Psychophysics* 71 (7), 1439–59.
LeDoux, J. E., and Damasio, A. R. (2013). "Emotions and Feelings," in Kandel et al. (eds). *Principles of Neural Science*, pp. 1079–94.
Lehrer, K. (1979). "The Gettier Problem and the Analysis of Knowledge," in G. Pappas (ed.), *Justification and Knowledge*. Dordrecht: D. Reidel, pp. 65–78.
Leibniz, G. W. (1996). *Leibniz: New Essays on Human Understanding*. Cambridge: Cambridge University Press.
Lewis, D. K. (1979). "Attitudes *De Dicto* and *De Se*," *Philosophical Review* 88: 513–43.
Levine, J. (2001). *Purple Haze*. Oxford: Oxford University Press.
Li, W., Howard J. D., Parrish, T. B., and Gottfried, J. A. (2008). "Aversive Learning Enhances Perceptual and Cortical Discrimination of Indiscriminable Odor Cues," *Science* 319: 1842–5.
Li, Z., Phillips, J., and Durgin, F. H. (2011). "The Underestimation of Egocentric Distance: Evidence from Frontal Matching Tasks," *Attention, Perception, & Psychophysics* 73(7): 2205–17.
Liu, M. (2021). "The Polysemy View of Pain," *Mind & Language*: DOI: 10.1111/mila.12389
Loevenbruck, H., Grandchamp, R., Rapin, L., Nalborczyk, L., Dohen, M., Perrier, P., Baciu, M., and Perrone-Bertolotti, M. (2018). "A Cognitive Neuroscience View of Inner Language: To Predict and to Hear, See, Feel," in Langland-Hassan and Vicente (eds.), *Inner Speech: New Voices*, pp. 131–67.
Logothetis, N. K., and Sheinberg, D. L. (1996). "Visual Object Recognition," *Annual Review of Neuroscience* 19: 577–621.
Loomis, J. M., and Philbeck, J. W. (2008). "Measuring Spatial Perception with Spatial Updating and Action," in R. L. Klatzky, B. MacWhinney, and M. Behrmann (eds.), *Embodiment, Ego-Space, and Action*. New York: Psychology Press, pp. 1–43.
Loomis, J. M., Da Silva, J. A., Philbeck, J. W., and Fukusima, S. S. (1996). "Visual Perception of Location and Distance," *Current Directions in Psychological Science* 5: 72–7.
Loomis, J. M., Klatzky, R. L., Philbeck, J. W., and Golledge, R. D. (1998). "Assessing Auditory Distance Perception Using Perceptually Directed Action," *Perception & Psychophysics* 60(6): 966–80.
Lupyan, G. (2009). "Extra-communicative Functions of Language: Verbal Interference Causes Selective Categorization Impairments," *Psychonomic Bulletin & Review* 16(4): 711–18.
Lycan, W. G. (1988). *Judgment and Justification*. Cambridge: Cambridge University Press.
Lycan, W. G. (2004). "The Superiority of HOT to HOP," in R. J. Gennaro (ed.), *Higher-Order Theories of Consciousness: An Anthology*. Philadelphia: John Benjamins, pp. 93–114.
Lyons, J. C. (2019). "Algorithm and Parameters: Solving the Generality Problem for Reliabilism," *Philosophical Review* 128(4): 463–509.
McBride, D. M., and Cutting, J. C. (2019). *Cognitive Science*, 2nd edition. Thousand Oaks, CA: Sage Publications.
McCain, K. (2014). *Evidentialism and Epistemic Justification*. New York: Routledge.
McCain, K., and Poston, T., eds. (2017). *Best Explanations: New Essays on Inference to the Best Explanation*. Oxford: Oxford University Press.
McDowell, J. (1982). "Criteria, Defeasibility, and Knowledge," *Proceedings of the British Academy* 68: 455–79.
McDowell, J. (1996). *Mind and World*. Cambridge, MA: Harvard University Press.
MacEvoy S., and Paradiso, M. (2001). "Lightness Constancy in Primary Visual Cortex," *Proceedings of the National Academy of Sciences* 98: 8827–31.

McGann, J. P. (2018). "Associative Learning and Sensory Neuroplasticity: How Does It Happen and What Is It Good For?" *Learning and Memory* 22: 567–76.
McGinn, C. (1982). *The Character of Mind*. Oxford: Oxford University Press.
Machery, E. (2009). *Doing Away with Concepts*. Cambridge: Cambridge University Press.
Mandelbaum, E. (2017). "Seeing and Conceptualizing: Modularity and the Shallow Contents of Perception," *Philosophy and Phenomenological Research*: 1–17.
Marlow, P. J., Kim, J., and Anderson, B. (2012). "The Perception and Misperception of Specular Surface Reflectance." *Current Biology* 22(20): 1909–13.
Marshall, J. C., and Halligan, P. W. (1988). "Blindsight and Insight in Visuo-spatial Neglect," *Nature* 336: 766–7.
Martin, M. G. F. (1998). "Setting Things before the Mind," *Royal Institute of Philosophy Supplement* 43: 157–79.
Matthen, M. (2021). "The Dual Structure of Touch," in F. de Vignemont, A. Serino, H. Y. Wong, and A. Farnè (eds.), *The World at Our Fingertips*. Oxford: Oxford University Press, pp. 197–214.
Meinong, A. (1981). "The Theory of Objects," in R. M. Chisholm (ed.), *Realism and the Background of Phenomenology*. Atascadero, CA: Ridgeview, pp. 76–117.
McDowell, J. (2013). "Tyler Burge on Disjunctivism (II)," *Philosophical Explorations* 16(3): 259–79.
Melzack, R., Wall, P. D., and Ty, T. C. (1982). "Acute Pain in an Emergency Clinic: Latency of Onset and Descriptor Patterns Related to Different Injuries," *Pain* 14: 33–43.
Miceli, G., Fouch, E., Capasso, R., Shelton, J., Tomaiuolo, F., and Caramazza, A. (2001). "The Dissociation of Color from Form and Function Knowledge," *Nature* 4(6): 662–7.
Middlebrooks, J. C., and Green, D. M. (1991). "Sound Localization by Human Listeners," *Annual Review of Psychology* 42: 135–59.
Mill, J. (1869). *Analysis of the Phenomena of the Human Mind*, 2nd edition, Volume I. London: Longman, Green, Reader & Dyer.
Millar, B. (2019). "Smelling Objects," *Synthese* 196: 4279–303.
Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.
Millikan, R. G. (1986). "Thoughts without Laws: Cognitive Science without Content," *Philosophical Review* 95: 47–80.
Millikan, R. G. (1989a). "Biosemantics," *Journal of Philosophy* 86: 281–97.
Millikan, R. G. (1989b). "In Defense of Proper Functions," *Philosophy of Science*. 56(2): 288–302.
Morales, J., Bax, A., and Firestone, C. (2020). "Sustained Representation of Perspectival Shape," *Proceedings of the National Academy of Sciences*, 117: 14873–82.
Morales, J. and Firestone, C. (forthcoming). "Empirical Evidence for Perspectival Similarity: Comment on Burge & Burge."
Moutoussis, K., and Zeki, S. (2002). "The Relationship between Cortical Activation and Perception Investigated with Invisible Stimuli," *Proceedings of the National Academy of Sciences* 99: 9527–32.
Murphy, G. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.
Murray, I. J., Parry, N. R. A., McKeefry, D. J., and Panorgias, A. (2012). "Sex-related Differences in Peripheral Human Color Vision: A Color Matching Study," *Journal of Vision* 12(1): 1–10.
Murray, S. O., Boyaci, H., and Kersten, D. (2006). "The Representation of Perceived Angular Size in Human Primary Visual Cortex," *Nature Neuroscience* 9: 429–34.

Nagel, T. (1979). *Mortal Questions*. Cambridge: Cambridge University Press.
Neander, K. (1991). "Functions as Selected Effects," *Philosophy of Science*, 58: 168–84.
Neander, K. (1996). "Swampman Meets Swampcow," *Mind and Language* 11: 118–29.
Neander, K. (2017). *A Mark of the Mental: In Defense of Informational Teleosemantics*.
Neitz, J., and Jacobs, G. H. (1986). "Polymorphism of the Long-wavelength Cone in Normal Human Colour Vision," *Nature* 323: 623–5.
Neitz J., Neitz, M., and Jacobs G. H. (1993). "More than Three Different Cone Pigments among People with Normal Color Vision," *Vision Research* 33: 117–22.
Newton A. M. and deVilliers, J. G. (2007). "Thinking While Talking: Adults Fail Non-verbal False-belief Reasoning," *Psychological Science* 18(7): 574–9.
Nguyen, D., Naffziger, E. E., and Berridge, K. C. (2021). "Positive Affect: Nature and Brain Bases of Liking and Wanting," *Current Opinion in Behavioral Sciences* 39: 72–8.
Nichols, S., and Stich, S. P. (2003). *Mindreading*. Oxford: Oxford University Press.
Noë, A. (2006). *Action in Perception*. Cambridge, MA: MIT Press.
Norman, J. F., Todd, J. T., Perotti, V. J., Tittle, J. S. (1996). "The Visual Perception of Three-dimensional Length," *Journal of Experimental Psychology: Human Perception and Performance*," 22: 173–86.
O'Callaghan, C. (2007). *Sounds: A Philosophical Theory*. Oxford: Oxford University Press.
O'Keefe, J., and Burgess, N. (1996). "Geometric Determinants of the Place Fields of Hippocampal Neurons," *Nature* 381: 425–8.
O'Keefe, J., and Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press.
Oliva, A. (2005). "Gist of the Scene," in L. Itti, G. Rees, and J. K. Tsotsos (eds.), *Encyclopedia of Neurobiology of Attention*. San Diego, CA: Elsevier, pp. 251–6.
Oliva, A., and Torralba, A. (2006). "Building the Gist of a Scene: The Role of Global Image Features in Recognition," *Progress in Brain Research* 155: 23–36.
Palmer, S. (1999). *Vision Science*. Cambridge, MA: MIT Press.
Papineau, D. (1987). *Reality and Representation*. Oxford: Blackwell.
Papineau, D. (2001). "The Status of Teleosemantics, or How to Stop Worrying about Swampman," *Australasian Journal of Philosophy* 79(2): 279–89.
Pappas, G. (2014). "Internalist vs. Externalist Conceptions of Epistemic Justification," *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/justep-intext/
Parma, V., Ferraro, S., Miller, S. S., Åhs, F., and Lundström, N. (2015). "Enhancement of Odor Sensitivity Following Repeated Odor and Visual Fear Conditioning," *Chemical Senses* 40: 497–506.
Parsons, T. (1980). *Nonexistent Objects*. New Haven: Yale University Press.
Pasnau, R. (1999). "What Is Sound?" *Philosophical Quarterly* 49: 309–24.
Pautz, A. (2010). "Do Theories of Consciousness Rest on a Mistake?" *Philosophical Issues* 20: 333–67.
Pautz, A. (2021). *Perception*. New York: Routledge.
Peacocke, C. (1984). *Sense and Content*. Oxford: Oxford University Press.
Peng, Y., Gillis-Smith, S., Jin, H., Tränkner, D., Ryba, N. J. P., and Zuker, C. S. (2015). "Sweet and Bitter Taste in the Brain of Awake Behaving Animals," *Nature* 527: 512–23.
Penrose, L. S., and Penrose, R. (1958). "Impossible Objects: A Special Type of Visual Illusion," *The British Journal of Psychology* 49(1): 31–3.
Pitts, M., Metzler, S., and Hillyard, S. (2014). "Isolating Neural Correlates of Conscious Perception from Neural Correlates of Reporting One's Perception," *Frontiers in Psychology*, 5(1078), 1–16.

Polley, D. B., Heiser, M. A., Blake, D. T., Schreiner, C. E., and Merzenich, M. M. (2004). "Associative Learning Shapes the Neural Code for Stimulus Magnitude in Primary Auditory Cortex," *Proceedings of the National Academy of Sciences* 101: 16351–6.

Pollock, J. (1984). "Reliability and Justified Belief", *Canadian Journal of Philosophy* 14: 103–14.

Pollock, J. (1986). *Contemporary Theories of Knowledge*. Totowa, NJ: Rowman and Littlefield.

Pooresmaeli, A., Arrighi, R., Biagi, L., and Marrone, M. C. (2013). "Blood Oxygen Level Dependent Activation of the Primary Visual Cortex Predicts Size Adaptation Illusion," *Journal of Neuroscience* 33(40): 15999–6008.

Porro, C. A., Cettolo, V., Fransescato, M. P., and Baraldi, P. (1998). "Temporal and Intensity Coding of Pain in Human Cortex," *Journal of Neurophysiology* 80(6): 3312–20.

Poston, T. (2014). *Reason & Explanation: A Defense of Explanatory Coherentism*. New York: Palgrave-Macmillan.

Price, D. D. (2000). "Psychological and Neural Mechanisms of the Affective Dimension of Pain," *Science* 288: 1769–72.

Price, D. D., and Bushnell, M. C., eds. (2004). *Psychological Methods of Pain Control: Basic Science and Clinical Perspectives*. Seattle: IASP Press.

Prinz, J. J. (2002). *Furnishing the Mind*. Cambridge, MA: MIT Press.

Prinz, J. J. (2012). *The Conscious Brain*. Oxford: Oxford University Press.

Pryor, J. (2000). "The Skeptic and the Dogmatist," *Noûs* 34 (4): 517–49.

Pryor, J. (2013). "Problems for Credulism," in C. Tucker (ed.), *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*. Oxford: Oxford University Press, pp. 89–133.

Putnam, H. (1975a). "Is Semantics Possible?" in Putnam, *Mind, Language, and Reality*. Cambridge: Cambridge University Press, pp. 139–52.

Putnam, H. (1975b). "The Meaning of 'Meaning'," in Putnam, *Mind, Language, and Reality* Cambridge: Cambridge University Press, pp. 215–71.

Pylyshyn, Z. W. (2003). *Seeing and Visualizing: It's Not What You Think*. Cambridge, MA: MIT Press.

Pylyshyn, Z. W. (2007). *Things and Places: How the Mind Connects with the World*. Cambridge, MA: MIT Press.

Quian Quiroga, R. (2017). "How Do We Recognize a Face?" *Cell* 159: 975–7.

Quian Quiroga, R., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). "Invariant Visual Representation by Single Neurons in the Human Brain," *Nature* 435: 1102–7.

Quilty-Dunn, J. (2016). "Iconicity and the Format of Perception," *Journal of Consciousness Studies* 23(3–4): 255–63.

Quine, W. V. (1960). *Word and Object*. Cambridge, MA: MIT Press.

Rainville, P. (2004). "Pain and Emotions," in Price and Bushnell (eds.), *Psychological Methods of Pain Control: Basic Science and Clinical Perspective*, pp. 117–41.

Ratliff, K., and Newcombe, N. (2008). "Is Language Necessary for Human Spatial Reorientation? Reconsidering Evidence from Dual Task Paradigms," *Cognitive Psychology* 56(2): 142–63.

Rescorla, M. (2018). "Maps in the Head?" in K. Andrews and J. Beck (eds.), *The Routledge Handbook of Philosophy of Animal Minds*. Milton Park, England: Routledge, pp. 34–45.

Rhudy, J. L., and Meagher, M. W. (2000). "Fear and Anxiety: Divergent Effects on Human Pain Thresholds," *Pain* 84: 65–75.

Rolls, E. T. (2019). *The Orbitofrontal Cortex*. Oxford: Oxford University Press.

Rosenthal, D. M. (2005). *Consciousness and Mind*. Oxford: Oxford University Press.

Russell, B. (1911). "Knowledge by Acquaintance and Knowledge by Description," *Proceedings of the Aristotelian Society* 11: 108–28.
Russell, B. (1912). *The Problems of Philosophy*, London: Williams and Norgate; New York: Henry Holt and Company.
Russell, B. (1914). "On the Nature of Acquaintance," *Monist* 24: 1–16, 161–87, 435–53.
Sandell, G. J., and Chronopoulos, M. (1997). "Perceptual Constancy of Musical Instrument Timbres: Generalizing Timbre Knowledge across Registers," *Proceedings of the Third Triennial ESCOM Conference*. Uppsala: Uppsala University.
Sapolsky, R. M. (2017). *Behave*. New York: Penguin Books.
Sartre, J. P. (1956). *Being and Nothingness*. New York: Philosophical Library.
Schellenberg, S. (2010). "The Particularity and Phenomenology of Perceptual Experience," *Philosophical Studies* 149: 19–48.
Schellenberg, S. (2016). "Perceptual Particularity," *Philosophy and Phenomenological Research* 93: 25–54.
Schellenberg, S. (2018). *The Unity of Perception: Content, Consciousness, Evidence*. Oxford: Oxford University Press.
Schilder, P., and Stengel, E. (1931). "Asymbolia for Pain," *Archives of Neurology & Psychiatry*, 25: 598–600.
Scholl, B. J. (2007). "Object Persistence in Philosophy and Psychology," *Mind and Language* 22: 563–91.
Schooler, J. W., and Maus, I. B. (2010). "To Be Happy and to Know It: The Experience and Meta-Awareness of Pleasure," in Kringelbach and Berridge (eds.), *Pleasures of the Brain*, pp. 244–54.
Schultz, W. (2006). "Behavioral Theories and the Neurophysiology of Reward," *Annual Review of Psychology* 57: 87–115.
Schultz, W., Dayan, P., and Montague, P. R. (1997). "A Neural Substrate of Prediction and Reward," *Science* 275: 1593–8.
Schultz, W., Tremblay, L., and Hollerman, J. (2000). "Reward Processing in Primate Orbitofrontal Cortex and Basal Ganglia," *Cerebral Cortex* 10: 272–83.
Schwitzgebel, E. (2008). "The Unreliability of Naive Introspection," *The Philosophical Review* 117: 245–73.
Scott, M. (2013). "Corollary Discharge Provides the Sensory Content of Inner Speech," *Psychological Science* 24: 1824–30.
Scott, M., Young, H. H., Gick, B., and Werker, J. (2013). "Inner Speech Captures the Perception of External Speech," *Journal of the Acoustic Society of America* 133.
Seligman, M. E. P., and Hager, J. L. (1972). "Biological Boundaries of Learning. The Sauce-bearnaise Syndrome," *Psychology Today* 6: 59–61, 84–7.
Shea, N. (2014). "Reward Prediction Error Signals are Meta-Representational," *Noûs* 48(2): 314–41.
Shea, N. (2018). *Representation in Cognitive Science*. Oxford: Oxford University Press.
Shepard, R. (1978). "The Mental Image," *American Psychologist* 33(2): 125–37.
Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C. R., Murray, R. M., and McGuire, P. K. (2001). "A Functional Study of Auditory Verbal Imagery," *Psychological Medicine* 31: 241–53.
Skelton, A., Catchpole, G., Abbott, J., Bosten, J., and Franklin, A. (2017). "Biological Origins of Color Categorization," *Proceedings of the National Academy of Sciences* 114(21), 5545–50.
Smith, E. E., and Medin, D. L. (1981). *Categories and Concepts*. Cambridge, MA: Harvard University Press.

Smith, K. S., Mahler, S. V., Pecina, S., and Berridge, K. C. (2010). "Hedonic Hotpots: Generating Sensory Pleasure in the Brain," in Kringelbach and Berridge (eds.) *Pleasures of the Brain*, pp. 27–49.

Smithies, D. (2019). *The Epistemic Significance of Consciousness*. Oxford: Oxford University Press.

Song, I., and Keil, A. (2014). "Differential Classical Conditioning Selectively Heightens Response Gain of Neural Population Activity in Human Visual Cortex," *Psychophysiology* 51: 1185–94.

Soteriou, M. (2000). "The Particularity of Visual Perception," *European Journal of Philosophy* 8: 173–89.

Spelke, E. S. (1988). "Where Perceiving Ends and Thinking Begins: The Apprehension of Objects in Infancy," in A. Yonas (ed.), *Perceptual Development in Infancy. Minnesota Symposium on Child Psychology* 20. Hillsdale, NJ: Erlbaum, pp. 191–210.

Spelke, E. S. (1990). "Principles of Object Perception," *Cognitive Science* 14: 29–56.

Spence, C., Levitan, C. A., Shankar, M. U., and Zampini, M. (2015). "Does Food Color Influence Taste and Flavor Perception in Humans?" *Chemosensory Perception* 3: 68–84.

Sperandio, I., Chouinard, P., and Goodale, M. (2012). "Retinotopic Activity in V1 Reflects the Perceived and Not the Retinal Size of an Afterimage," *Nature Neuroscience* 15: 540–2.

Sperling, G. (1960). "The Information Available in Brief Visual Presentations," *Psychological Monographs* 74: 1–29.

Stampe, D. W. (1977). "Toward a Causal Theory of Linguistic Representation," *Midwest Studies in Philosophy* 2(1): 42–63.

Starr, C. J., Sawaki, L., Wittenberg, G. F., Burdette, J. H., Oshiro, Y., Quevedo, A. S., and Coghill, R. C. (2009). "Roles of the Insular Cortex in the Modulation of Pain: Insights from Brain Lesions," *Journal of Neuroscience* 29(9): 2684–94.

Stevens, S. S. (1962). "The Surprising Simplicity of Sensory Metrics," *American Psychologist*, 17(1): 29–39.

Tarr, M. J., and Bülthoff, H. H. (1995). "Is Human Object Recognition Better Described by Geon Structural Descriptions or by Multiple Views? Comment on Biederman and Gerhardstein," *Journal of Experimental Psychology: Human Perception & Performance* 21(6): 1494–505.

Tarr, M. J., Williams, P., Hayward, W. G., and Gauthier, I. (1998). "Three-dimensional Object Recognition is Viewpoint-dependent," *Nature Neuroscience* 1: 275–7.

Taylor-Clarke, M., Jacobsen, P., Haggard, P. (2004). "Keeping the World a Constant Size: Object Constancy in Human Touch," *Nature Neuroscience* 7: 219–20.

Thouless, R. H. (1931a). "Phenomenal Regression to the Real Object I," *Journal of Psychology* 20: 339–59.

Thouless, R. H. (1931b). "Phenomenal Regression to the Real Object II," *Journal of Psychology* 21: 20–30.

Thouless, R. H. (1972). "Perceptual Constancy or Perceptual Compromise," *Australian Journal of Psychology* 24: 133–40.

Tian, X., and Poeppel, D. (2010). "Mental Imagery of Speech and Movement Implicates the Dynamics of Internal forward Models," *Frontiers in Psychology* 1, #166, doi.org/10.3389/fpsyg.2010.00166.

Tian, X., and Poeppel, D. (2013). "The Effect of Imagination on Stimulation: The Functional Specificity of Efference Copies in Speech Processing," *Journal of Cognitive Neuroscience* 25: 1020–36.

Tiest, W. M. B., and Kappers, A. M. L. (2007). "Haptic and Visual Perception of Roughness," *Acta Psychologica*, 124(2): 177–89.

Tittle, J. S., Todd, J. T., Perotti, V. J., and Norman, J. F. (1995). "Systematic Distortion of Perceived Three-Dimensional Structure from Motion and Binocular Stereopsis," *Journal of Experimental Psychology*: *Human Perception and Performance* 21(3): 663–78.

Todd, J. T., Oomes, A. H. J., Koenderink, J. J., and Kappers, A. M. L. (2001). "On the Affine Structure of Perceptual Space," *Psychological Science* 12(3): 191–6.

Tsao, D. Y. (2019). "Face Values," *Scientific American* 320: 22–9.

Tye, M. (2002). *Consciousness, Color and Content*. Cambridge, MA: MIT Press.

Tye, M. (2009). *Consciousness Revisited: Materialism Without Phenomenal Concepts*. Cambridge, MA: MIT Press.

Tye, M. (2014). "What is the Content of a Hallucinatory Experience?" in B. Brogaard (ed.), *Does Perception Have Content?* Oxford: Oxford University Press, pp. 291–308.

van Zandvoort, M. J. E., Nijboer, T. C. W., and de Haan, E. (2007). "Developmental Colour Agnosia," *Cortex* 43(6): 750–7.

Vineberg, S. (2016). "Dutch Books," *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/dutch-book/.

Vogel, J. (1990). "Cartesian Skepticism and Inference to the Best Explanation," *The Journal of Philosophy* 87(11): 658–66.

von der Malsburg, C. (1999). "The What and Why of Binding: The Modeler's Perspective," *Neuron* 24(1): 95–104.

Wagner, M. (2006). *The Geometry of Visual Space*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ward, E. J., and Scholl, B. J. (2015). "Inattentional Blindness Reflects Limitations on Perception, not Memory: Evidence from Repeated Failures of Awareness," *Psychonomic Bulletin & Review* 22(3): 694–700.

Ward, E. J., Bear, A., and Scholl, B. J. (2016). "Can You Perceive Ensembles without Perceiving Individuals? The Role of Statistical Perception in Determining whether Awareness Overflows Access," *Cognition* 152: 78–86.

Warren, W. H. (2005). "Direct Perception: The View from Here," *Philosophical Topics* 33(1): 335–61.

Warren, W. H. (2021). "Information is Where You Find It: Perception as an Ecologically Well-posed Problem," *i-Perception* 12(2): 1–24.

Warrier, C. M., and Zatorre, R. J. (2004). "Right Temporal Cortex is Critical for Utilization of Melodic Temporal Cues in a Pitch Constancy Task," *Brain* 127: 1616–25.

Weisberg, J. (2012). "The Bootstrapping Problem," *Philosophy Compass* 7(9): 597–610.

Wiggins, D. (2001). *Sameness and Substance Renewed*. Cambridge: Cambridge University Press.

Williams, J. R. G. (2019). *The Metaphysics of Representation*. Oxford: Oxford University Press.

Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., and Boroditsky, L. (2007). "Russian Blues Reveal Effects of Language on Color Discrimination," *Proceedings of the National Academy of Sciences* 104(19): 7780–5.

Wittgenstein, L. (1953). *Philosophical Investigations*. New York: Macmillan.

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. London: Routledge and Kegan Paul.

Wolfe, J. M., Kluender, K. R., and Levi, D. M. (2006). *Sensation and Perception*. Sunderland, MA: Sinauer Associates.

Wouter M. B., Tiest, W. M. B., and Kappers, A. M. L (2007). "Haptic and Visual Perception of Roughness," *Acta Psychologica* 124: 177–89.

Wu, W. (2022). *Movements of the Mind*. Oxford: Oxford University Press.

Yantis, S. (2014). *Sensation and Perception*. New York: Worth Publishers.

Yeshuran, Y., and Carrasco, M. (1999). "Attention Improves Performance in Spatial Resolution Tasks," *Vision Research* 39: 293–30.

Yoshioka, T., Craig, J. C., Beck, G. C., and Hsiao, S. S. (2011). "Perceptual Constancy of Texture Roughness in the Tactile System," *The Journal of Neuroscience* 31(48): 17603–11.

Zahoric, P., and Wightman F. L. (2001). "Loudness Constancy with Varying Sound Source Distance," *Nature Neuroscience* 4(1): 78–83.

Zhang, Y., Zhao, Y., Zhu, X., Sun, X., and Zhou, X. (2013). "Refining Cortical Representation of Sound Azimuths by Auditory Discrimination Training," *Journal of Neuroscience* 33: 9693–8.

Index

Note: Figures are indicated by an italic "*f*" following the page numbers.

For the benefit of digital users, indexed terms that span two pages (e.g., 52–53) may, on occasion, appear on only one of those pages.