Mindmelding

Consciousness, Neuroscience, and the Mind's Privacy

William Hirstein
Dept. of Philosophy,
Elmhurst College
Elmhurst, Illinois, USA

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

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 in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2012

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Library of Congress Control Number: 2011944055

Typeset in Minion by Cenveo, Bangalore, India
Printed in Great Britain
on acid-free paper by
CPI Group (UK) Ltd, Croydon, CR0 4YY

ISBN 978–0–19–923190–4

10 9 8 7 6 5 4 3 2 1

Preface

Somehow a surprising number of philosophers and scientists have got themselves into the position of claiming that the conscious mind can never be explained in straightforwardly physical terms. The main sticking point concerns the mind's fundamental privacy, they believe. Since the features of our conscious states can only ever be experienced by one person, even if they are ultimately physical, they seem to exist in a different class of physical things from normal objects such as rocks or chairs. The main difference between that special inside world of consciousness and the external world concerns the way we gain knowledge about them. The difference between the way our thoughts appear from the inside, to us, and the way they appear to the outside, to the scientific observer, is just too great, they argue. When I look at that red flower there, I might actually see it as green due to a form of color blindness. But there is no way that scientists or anyone else could know that by studying me from the outside, they say. The scientists won't see anything green *or* red in my brain, just lots of gray and white matter. I disagree, and in the following pages, I will sketch an alternative framework for understanding the mind and consciousness in straightforward physical terms. I will combine information from neuroscience, psychology, and philosophy to construct this alternative picture. One interesting and testable consequence of this view is the possibility of something I call "mindmelding," allowing one person to directly experience the conscious states of another.

One message I have for scientists is that there are clear and commonsensical philosophical positions one can take on the problems associated with mind. No extreme philosophical acrobatics, new metaphysical categories, paradigm shifts, or supreme mental acts of comprehension are required to understand our conscious minds in physical terms. We can even use our existing concepts of the mind (including *thought*, *consciousness*, *awareness*, *belief*, *desire*, *emotion*, *memory*) as we come to understand the brain.

This book is for anyone interested in the mind and brain. It will be of special interest to people who have thought about the mind–body problem, now known as the problem of consciousness. Within that group, students, teachers, and researchers in the cognitive sciences should find much of interest here. Among the disciplines that make up cognitive science (psychology, philosophy, artificial intelligence, neuroscience, linguistics, and anthropology), I am speaking primarily to neuroscientists, psychologists, to those in the new hybrid field of cognitive neuropsychology, and to philosophers. This book contains an introduction to the neuroscience of consciousness that should benefit philosophers. It also approaches the philosophical issues surrounding consciousness in ways that I hope will be accessible to scientists of all sorts, with the caveat noted above that the views here differ significantly from the philosophical mainstream.

I would like to thank Patrick Ackles, John Bickle, William Brenner, Melinda Campbell, Todd Feinberg, Brian Keeley, Erik Myin, Thomas Natsoulas, V. S. Ramachandran, Nadia Sahely, John R. Searle, and Katrina Sifferd. I would also like to thank these students, who brought these issues to life in class and in discussions outside of class with their creativity and critical insight: Xavier Arko, Katherine Banks, Sandra Bluhm, Eduardo Cordon, Andrew Hartley, Matthew Kovich, Kevin Lepore, David Metcalfe, Joseph Nicolini, Sarah Roe, Robert Tempio, and Vincent Thomas.

Contents

Chapter 1

The impasse

Is the mind private?

Close your eyes for a moment, and form a mental image. Make it of something definite; perhaps try imagining your mother's face. Got it? Now, what *is* that image? What is it made out of? Is it physical, or non-physical? Is the mind just a particularly lively form of matter, the flashing of electrical impulses in the brain? Or is it made of something much finer, something non-physical? The mystery of how to answer this question has traditionally been known as the mind–body problem. Now it is more frequently referred to as the problem of consciousness. A majority of scientists and philosophers today, myself included, believe that the mind exists as a set of physical brain states and events. This makes them *materialists*. Materialism is a philosophical/scientific theory according to which everything is physical, including the mind. Everything that happens in the mind—thoughts, images, and feelings—is actually a physical state or event. Rapid progress in the neurosciences is forcing our ancient questions about the connection between mind and brain into the light. Given what we know about the anatomy of the brain at many different levels, if the mind and its consciousness are somewhere in there, we ought to be able to say where. After ignoring consciousness for several decades, scientists have lately accepted the challenge of answering in detail all the questions about where and how the brain produces conscious states (for some examples, see Crick and Koch, 1994; Damasio, 1999; Rolls, 1999; and Llinás, 2002).

There is a lot more at stake here than the considered opinions of a bunch of scientists and philosophers, of course. There are competing conceptions of the way we understand our minds, of free will and our notions of responsibility for what we do, and of who we really are, behind each of the positions. Some of these conceptions threaten to force wholesale changes in the ways we conceive of ourselves. The materialists can't guarantee that as we unearth the neural mechanisms which produce the mind, we won't learn things that are destructive of the ideas we have of ourselves. Materialists need to be certain they are right before they put people through all the adjustments in beliefs that their success might require. They need to be certain that they have found theories powerful and detailed enough to construct all the explanations required. History has seen their claims before. Materialism has existed for thousands of years. In the West, it is traceable to the ancient Greek atomists, Democritus and Leucippus, who believed that the soul was constructed of very fine atoms of matter.

Who exactly are we, and what is our mission here on earth? Besides questions about how the universe began, the early Greeks were especially interested in the question of

what our "excellence" is—which they called our *arête*. The excellence of birds is to fly, horses to run. What is the function we perform best; the one we were meant to do, or the one it is best for us to do? If we understand the nature of our minds, we will be in a better position to answer this question. If the human mind is not physical, there is an entire realm of human life that no other species possesses, this makes us seem special, or even hints at a special creation. But a materialist theory of the mind might also portray us as special. We might be special in that the brain events that embody our minds are unique to us as a species on earth. The actual physical events that constitute our minds are astonishingly complex and subtle. The world we make must be based on our knowledge of what we are good at and bad at, of what we are capable of doing, and incapable of doing. It must be based firmly on our knowledge of ourselves. If we want to know how to raise children to behave ethically, as well as how to cause adults to act more ethically, we need to understand the sources of weakness we have that lead to unethical behavior. The brain is not a topic-neutral computing device, like a Turing machine. It is specifically designed to work with certain types of input and produce certain types of output. Certain kinds of computations come naturally to the brain, others we find difficult and unintuitive.

Materialism has always been blocked by its inability to accomplish the first thing on its list of desiderata: to explain exactly how mere matter can account for the amazing things the mind can do. This allowed dualism, the view that the mind is non-physical, to reign as the majority view among philosophers and scientists until the middle of the 20th century, when the advent of the computer age began to convince many people that mental processes could be accomplished by physical machines. But computers, despite their early promise, have so far proved to be unable to equal the creativity and flexibility of the human mind. In the early 1980s, however, a revolution began quietly boiling away in the biological neurosciences, driven by powerful new experimental techniques that allowed the monitoring of mental events as they occur, via brain imaging. At the same time, our understanding of brain anatomy and physiology began expanding rapidly, due to powerful new investigation techniques, such as cell-staining studies, as well as burgeoning research in neurochemistry and neurocomputation.

Working inward from the sense organs, neuroscientists have begun to isolate those brain areas and processes important for consciousness. After starting at the back of the brain, where most perceptual processing takes place, we have now made it all the way to the front. We know the basic wiring of the brain, and for the first time, we have theories of how it works as a whole. Yet a troubling problem remains. The ability to actually see the brain performing its different cognitive tasks, together with the converging data from neurophysiology and neuroanatomy, has led to the theory that conscious brain states involve large parts of the brain's cortex—its wrinkled outer covering—supported by activity in many subcortical structures, working together in synchronized rhythms. We now have our first detailed neurophysiological theories of consciousness. These theories are still operating from the outside, however. All of the current research techniques leave the scientific observer of the brain locked out of the experience the subject herself is having. Using their new imaging technologies, the scientists can observe all sorts of brain

activity, but it seems they can never detect the most crucial properties of conscious states, the ones we are aware of. If someone is looking at a blue sky, for instance, the scientists monitoring her brain can't detect anything blue. How, then, can we ever be certain that these synchronized brain states are the conscious mental events we experience? The possibility still seems entirely open that our conscious mental lives reside elsewhere, in some other realm.

This describes the current state of research, but can the scientists of the future ever enter that sacred citadel, the mind itself? If they can't, if conscious states are necessarily and permanently private, this creates a fundamental problem for materialism. If mental states lack a basic feature possessed by all known physical states—the capacity to be observed or experienced by more than one person—this is reason to suspect that they are not truly physical. We can put this in the form of an argument:

The privacy argument
Premise 1: No physical states are private.
Premise 2: All conscious states are private.

Conclusion: No conscious states are physical states.

Dualists tend to agree with both the premises and conclusion of this argument, since they have held all along that the mind is not physical. Materialists will obviously disagree with the conclusion, but the argument itself is formally valid, meaning that the conclusion follows logically from the premises. This means that any materialist who disputes the conclusion must find something wrong with one or both of the premises. There are several schools of thought among the materialists as to what the problem is.

One group of materialists accepts that conscious states are private, but argues that they are nevertheless physical states. Hence they reject the first premise of the privacy argument. But if there is a kind of physical state that is private, this sharply distinguishes them from the other physical states. This has led some materialists to posit the existence of a previously unknown category of entities, or a new *metaphysical category*. According to John Searle, their most prominent member, there is a special category of private physical states that includes our conscious states. This view has also been embraced by some scientists (see Damasio, 1999; Feinberg, 2001). One dramatic correlate of this view is the idea that we cannot separate the existence of conscious states from our knowledge or awareness of them. Unlike normal external objects, which exist whether we are aware of them or not, conscious states exist only when we are aware of them. In philosophical terms, the *ontology* of conscious states is necessarily bound to their *epistemology*. For them, to be is to be perceived; to exist is to be an object of someone's awareness.

The medieval philosopher William of Ockham warned against the creation of new such metaphysical categories, however, and his maxim, "Do not multiply entities beyond necessity," might be invoked here. There is always room for the revisionists such as Searle to argue, however, that in this case the new category of entities is necessary, in order to explain the mystery of consciousness. Surely though, there is force to what Ockham said: Given that the world is hard enough to understand, there needs to be some very clear

positive reason for creating new categories of things. It should not be something that we find ourselves backed into merely because we cannot figure something out.

Perhaps good communication is the way around the wall of privacy, a way to dispute the second premise of the privacy argument, the claim that all conscious states are private. According to Daniel Dennett, a combination of investigation from the outside and verbal reports of subjects is sufficient for us to gain all the knowledge we need about conscious states, so that there is nothing significant left over that is truly private. We can interview subjects at great length, posing questions to them to make sure we have understood what they are saying (Dennett, 1991). But, despite our talent for communication, there is a huge difference between actually having these experiences and hearing verbal reports about them. Anyone who ever tried to describe a dream to someone can attest to this. Verbal reports, no matter how thorough, can still be inadequate, misleading, or simply mistaken. The person hearing the verbal report is even further away from the conscious state than the scientist observing the functioning brain. Dennett's notion that the richness and subtlety of conscious experience can be captured in language also strikes one as deflationary. Verbal communication may work for thoughts that occur in linguistic form, but much of our mental lives consists of images and emotions.

Another very different type of approach that seems to be growing in popularity is that we will never be able to solve the dispute between materialists and dualists because we will never understand consciousness. We humans have finally met our intellectual match. According to this group, known as *mysterians,* the problem of relating consciousness to the physical realm cannot be solved, at least not by us (Chomsky, 1994; Nagel, 1998; McGinn, 1999). Colin McGinn, their primary spokesman, provides an interesting argument by analogy for mysterianism: Humans are biological organisms, formed by the processes of evolution. The other such species have limits to their mental capacities that are obvious to us. Dogs could never understand calculus, for instance. There must be, then, problems that are as far beyond our mental grasp as calculus is to that of dogs. We should be realistic and humble and admit this. This type of argument is not new of course, medieval philosophers argued that we could never understand God. It is disappointing, though, that something so close to us should prove ultimately unfathomable. Normally when one thinks of the limits of human understanding, one thinks of questions about God's nature or existence, or questions about distant events, in space or time. Will we ever know whether there is life on other planets? Will we ever understand how the universe began? Or events that lie beyond the limits of our perception? Can we solve the riddles of physics? Is there a limit to the universe? But given that neuroscientists are just beginning to understand the higher levels of brain function, and making extremely rapid progress, it seems absurd to give up now. The mind's problem-solving capacity has proven to be quite impressive thus far, and there is no sign of a slowing in the rate of its accomplishments. In fact, events seem more to be proceeding in the other direction: there are more sub-fields, more scientific journals, computers and communications get faster. If at some point well in the future, when neuroscientists are satisfied that they fully understand the brain, consciousness is still intractable, then perhaps we should start to worry.

The possibility of mindmelding

I am also going to argue that the second premise is false, but not in the deflationary way that Dennett does. Dennett lets us know about another person's conscious states by limiting what there is to know, not by expanding our abilities to know. Contrary to this, I think that our knowledge can be expanded in this realm. The way around the impasse is to question the cherished and little-examined assumption that one person can never directly experience the mind of another, what I call "mindmelding" (Hirstein, 2008). We are beginning to understand where in the brain the different constituents of conscious states reside, as well as how the brain knits them together to form the unified, coherent mental events we experience. I believe that this opens up the possibility of theorizing about how we might connect two brains to allow one person to directly experience the consciousness of another. That is the goal of this book, to show that the impasse is apparent, not real.

The possibility of one being having direct knowledge of the consciousness of another is already familiar to us. Many of us were brought up with the belief that God knows everything that happens in our minds. God presumably has to know this in order to evaluate our acts, since he needs to know what our options were as we perceived them. There are also science fiction tales about this happening, from the "Vulcan mindmeld" of the original *Star Trek*, where Mr Spock merges his consciousness with his subject, to the movie *Strange Days*, which depicts a device that can record one person's conscious experiences and allow another person to re-experience them by putting on a special headset and replaying the recording, to the current movie *Inception*, in which people mentally eavesdrop on the dreams of others. It seems at least *conceivable* that I could experience your consciousness. But how exactly can a normal present-day person, a mere human, have direct knowledge of what is happening in the mind of another? Is this genuinely possible, or is it one of those scenarios that merely seems possible but actually isn't, such as perpetual motion, or travel faster than the speed of light?

Another sort of objection to the idea of linking brains—to what one might call interpersonal introspection—is doubtless on the minds of many readers at this point. One idea that is at the heart of many thinkers' acceptance of privacy runs as follows: In order for one person to experience the consciousness of another, there needs to be a distinction in our minds between the experiencer and the object of experience. But this distinction cannot be made; there is no such thing as an independent experiencer in the mind, the objection goes. The problem is that it seems that both the state and our experience of it are combined in a simple, unbreakable, monolithic unit, as in Searle's view. A separate experiencer in the mind would be a *self*, and unlike the philosophers of old who wrote frequently about the self, many of today's writers take a more jaundiced view toward the idea. This view is supported by the failure of brain science to find a self, and some scientists have specifically denied the existence of a self in the brain (e.g., Metzinger, 2009). In the philosophical tradition, Hume (1739/1987) disputed that introspection reveals a self, and this tradition of self-skepticism continues to this day, through Wittgenstein (1955) to current writers such as Dennett (1991), who argues that the self is a sort of fiction.

Contrary to all these lines of thought, I believe that a clear separation can be made in the mind and brain between our conscious mental representations and other processes that both interact with these representations within consciousness and give rise to our sense of self, the sense of an experiencer. These two components are generated by different brain processes; hence the separation needed to allow us to explore the possibility of one person experiencing the conscious mental states of another person exists. Or so I hope to convince you. Showing that conscious states can be experienced by more than one person does not by itself show that they are physical, but it removes a huge block to thinking of them in that way. The idea that conscious states could in some real sense be shared by two people also allows us to avoid creating new metaphysical categories. It does not commit us to a new type of physical property, possessed by nervous systems, or on some accounts, by everything (see, e.g., Chalmers, 1996).

Assertions of privacy

Almost all of today's writers on consciousness accept the privacy—or what some refer to as the *subjectivity*—of mental states. Examples of this go back at least as far as John Locke, who said that we can never truly know what is in the mind of another, "because one man's mind could not pass into another man's body, to perceive what appearances were produced" (1690/1988). A collection of the current defenders of privacy includes a large portion of the current writers on consciousness, as the following quotations show:

> Current studies in brain imaging show that different parts of the brain are more or less active in different kinds of cognitive tasks, but the color patches on the brain maps can't show what the thoughts are, so we have to accept the identity [of "neural and mental events"] as an untestable theory.
>
> (Freeman, 2000, p.4)

> Now we psychologists are back studying subjective experiences: perceptions, recollections, intentions. But the problem remains: the mental things that we study have a completely different status from the material things that other scientists study. The only way I can know about the things in your mind is because you tell me about them. You press a button to tell me when you see the red light. But there is no way I can get into your mind and check the redness of your experience.
>
> (Frith, 2007, p.6)

> We think consciousness has to be largely private. By "private" we mean that it is accessible exclusively to the owner of the brain; it is impossible for me to convey to you the exact nature of my conscious percept of the color red, though I can convey information about it, such as whether two shades of red appear to me to be the same or different.
>
> (Crick and Koch, 2003, p.119)

> There appears to be a behaviorally defined brick wall, which I will call the subjectivity barrier, that limits which aspects of our conscious experience we can share and which we cannot, no matter how hard we try.
>
> (Palmer, 2003, p.195)

> You cannot have my experience in any direct fashion because of the subjectivity barrier.
>
> (Palmer, 2003, p.200)

Consciousness is an "entirely private, first-person phenomenon which occurs as part of the private first-person process we call mind" (Damasio, 1999, p.12). Damasio distinguishes between emotions, which can be measured objectively, and feelings, which are "always hidden, like all mental images necessarily are, unseen to anyone other than their rightful owner, the most private property of the organism is whose brain they occur" (Damasio, 2003, p.28).

Conscious states "are directly experienced only by single individuals." And, "what is directly experienced as qualia by one individual cannot be fully shared by another individual as an observer" (Edelman, 1994, p.114).

> There is something special about consciousness. Conscious experience arises as the workings of each individual brain. It cannot be shared under direct observation, as the physicist's objects can be shared. Thus, studying consciousness presents us with a curious dilemma: Introspection alone is not scientifically satisfactory, and though people's reports about their own consciousness are useful, they cannot reveal the workings of the brain underlying them. Yet, studies of the brain proper cannot, in themselves, convey what it is like to be conscious.
>
> (Edelman and Tononi, 2000, p.xi)

Having mental images is "a quintessential private event" (Kosslyn and Ochsner, 1994, p.290).

> Conscious experience, as an awareness of some thing or event, is directly accessible only to the individual having that experience, not to an external observer.
>
> (Libet, 1993, p.368)

Block reasons that, since the only way we can know about a person's conscious states is from his reports, "there seems to be a fundamental epistemic limitation in our ability to get a complete empirical theory of phenomenal consciousness" (Block, 2007b, p.483).

> Though I can easily observe another person, I cannot observe his or her subjectivity.
>
> (Searle, 1992, p.99)

> Every subjective phenomenon is essentially connected with a single point of view, and it seems inevitable that an objective, physical theory will abandon that point of view.
>
> (Nagel, 1974, p.437)

> The subjectivity of consciousness is an irreducible feature of reality.
>
> (Nagel, 1986, p.7)

> If your friend is staring at something green, you cannot look at her and see the ... this is not just an accidental fact; consciousness is necessarily not perceptible.
>
> (McGinn, 1999, p.114)

No instrument can breach the wall of privacy

Some writers are led from the observation that nothing can breach the wall of privacy to the claim that therefore the mind is not real:

> Of course, special instruments have been developed to help us make accurate measurements when we are dealing with very rare elements like radium, very small things like the base pairs in gene sequences, or very fast things like light. But these special instruments are, like magnifying glasses, simply extensions of your senses. They help us to see what is really there. No such instrument can help us see what is happening in the mind. The contents of the mind are not real.
>
> (Frith, 2007, p.16)

> I can see the edges of the universe with a telescope and I can see the activity in your brain with a scanner, but I can't "see" into your mind. The mental world, we all believe, is quite distinct from physical reality.
>
> (Frith, 2007, p.16)

Feinberg echoes this sentiment:

> There is no materiality to the experience 'pain' from the observer's point of view because the experience of pain from the inner point of view only exists as neural activity from the outside perspective.
>
> (Feinberg, 2001, p.146)

According to McGinn, this situation is permanent, not temporary:

> There is no way to modify or extend introspection and perception so that they can transcend their present limitations. That is like hoping that if we tinker with our sense of touch it will eventually give us perceptions of color.
>
> (McGinn, 1999, p.51)

> We cannot even conceive of a type of sense organ that would enable us to perceive consciousness... .
>
> (McGinn, 1999, p.114)

A method for directly determining the contents of someone's conscious state "apparently does not exist" laments Block (2007b, p.483).

Ten popular (but false) claims

The mind is primarily the conscious part of the brain. The brain produces, or generates the mind. To put it another way, the mind is a set of properties that the brain possesses, the main one being consciousness itself. It is wrong to say that the brain equals the mind, since there are all sorts of facts about the brain that are only distantly related to consciousness at best, and hence have little to do with the mind. There are all sorts of tricks that the brain uses to coordinate muscles, for instance, that don't concern our minds particularly. But neither is mind to be equated with consciousness. The mind certainly has an unconscious portion, you have all sorts of beliefs and memories that are not now conscious. But they exist, they are part of your mind. There could be an extreme theory that equated mind with consciousness, but it would be forever saddled with inadequate explanations of where our conscious thoughts, images, and feeling come from. The typical materialist theory of the brain contains explanations of how the unconscious becomes conscious, rather than treating conscious thought as arising *ex nihilo*, out of nothing. If you postulate that those things or properties arise out of nothing, you are violating another even more ancient and deeply held philosophical law: *Ex nihilo, nihilo est.* Or, nothing comes from nothing, i.e., everything has to come from something. Of course, this may be just another one of those limitations in our imaginations rather than in the world itself. We might be guilty of mistaking a psychological limitation for a metaphysical limitation.

In this section, I will describe a set of views, including privacy, each of which is widely held among today's philosophers and scientists. Most contemporary philosophers of mind hold most of these views, and if one overlaid all of the views of the current mainstream

writers in the philosophy of mind, as well as the scientific writers on the problem of consciousness, these views would stand out as areas of consensus. Philosophers will be surprised at the theorists grouped together here, since they are normally seen as being opposed to one another. It is true that they disagree about certain of the popular views, and most of their efforts are directed at one another on these points. They are all vying on the same field, however, and my contention is that they are on the wrong field. The list below represents those doctrines that are currently the most popular in the literature and among its consumers, i.e., practicing philosophers and scientists, graduate students, professionals in related fields, and a growing public audience. The internal consistency of these views, along with what are thought to be crushing arguments against any alternative view, suffice to account for the long run of popularity they have enjoyed.

Theories are designed to answer questions, dozens, hundreds, and thousands of them ultimately. These ten statements involve some of the most pressing questions facing theorists of consciousness today. As science and philosophy make progress this list will, of course, change. It will also grow, in that each of these questions will give rise to many finer-grained questions, some of which appear in what follows. These questions must be answered with a unified account, however. It will do no good to answer each one with a different theory, because one of the primary purposes of a theory is to unify our understanding of something, in a way that is clear to us and easy to apply and test. The right theory has to answer each question in a clear, concise, and convincing way.

Let me briefly describe these views, and equally briefly state why I believe they are false.

1. There is no way to separate the existence of conscious states from our knowledge or awareness of them. All of the other facts we know about have their own existence independent of our knowledge. I know that the Eiffel Tower is across the Atlantic, in Paris. The fact that the Eiffel Tower stands there in Paris in no way depends on my knowledge of it. The ordinary separation between our knowledge of a thing and its existence cannot be maintained in the case of conscious states. According to this view, conscious states exist only when someone is aware of them.

In Chapter 11, I will examine this claim, which I call the thesis of inseparability, in detail. I will argue there that the existence of our conscious states in fact can be separated from our knowledge of them. The alternative view saves the common-sense notion that we can always distinguish between our knowledge of a fact or thing, and the fact or thing itself (Hirstein, 2001, 2004). In the alternative framework I will develop, nothing has its existence necessarily tied to our knowledge of it. Thus there is no need to invent a new metaphysical category.

One way to question this claim is to examine our awareness of the visual periphery. We are aware of only a portion of our visual fields at any given time. One's focus of attention moves around within the visual field. It is typically located where the eyes are focused, but it need not be; one can visually attend to something without looking at it. We *can* be aware of the periphery outside the visual focus, but we normally are not. The unattended portions of the visual field are still part of a conscious brain state, just not one that we are currently aware of. One reason that it is wrong to say that we are normally aware of the periphery of our visual field is that there are now several experiments that show that the

information in the non-attended portions of the visual field cannot be reported by subjects (Mack and Rock, 1998; Rensink et al., 2003). This difference is captured by our everyday distinction between consciousness and consciousness of. Something can be conscious in our minds without our being conscious of it. Consciousness is necessary in order for someone to know about, or be aware of something, but it alone is not sufficient. The proper causal relations must exist between our conscious representations of that thing and a set of processes located primarily in the front of the brain called *executive processes*. The primary causal relation in this regard is what we normally call attention. Attention is an executive process and occurs when certain frontal processes engage in causal interaction with conscious representations existing in another part of the brain.

2. Conscious states are necessarily private. I will call this claim the thesis of privacy. Your conscious states are private in the sense that no one else will ever be able to know them in the intimate way that you do. Because of the nature of these states, it is impossible for another person to directly experience them. The defenders of privacy are not claiming that there are enormous technical and technological obstacles to knowing another's conscious states. There are problems we know of like this, such as the problem of "solving" chess, i.e., writing a computer program that can consider every possible move, or the problem of how humans can travel great distances in space. The popular view is that it is simply impossible for one person to know another's mind.

Contrary to this, I will argue that conscious states are normally, but not necessarily, private. In the alternative framework, the claim that one and only one person can know about conscious states is false. Our knowledge of our mental states is not an unbreakable simple event. It has a structure: Conscious representations interact with executive processes, and it is possible for these two components to reside in two different brains.

3. Some brain states and processes possess two mutually exclusive sets of properties, and we have different ways of knowing each. Imagine someone watching a parade. If our current knowledge of the neuroscience of human vision science is correct, her temporal lobes contain high-level visual processes, some of which play a role in embodying her conscious awareness of what she sees. According to several defenders of privacy, these parts of the brain have two different types of properties, conscious properties, and non-conscious properties. The person watching the parade is aware of the conscious properties in a direct way. On the other hand, brain scientists analyzing her brain would be aware of different, non-conscious properties, such as neural firing rates, quantities of neurotransmitters, and so on. The scientists are aware of these properties in a different way from the way in which the person herself is aware of the conscious properties.

I will argue instead that there is no essential metaphysical or epistemic distinction in the properties that brain states have. While the conscious properties are surprising and amazing, they are physical properties like the others, according to my alternative account. Brain states are physical states, and the properties they have are all physical properties. Every normal person can know about the instances of these properties in the most direct ways possible. There are more natural ways of knowing about these properties, such as

normal introspection, but we are not restricted to these. The alleged wall between my way of knowing about my conscious states by introspection, and others' ways of knowing, including scientists', can be breached from either direction. One person can know about the conscious states of another using something like interpersonal introspection, i.e., mindmelding. Or, we can attain knowledge of another's mental states using epistemic processes that are normally restricted to external ways of knowing, such as vision, by devising means to project mental images onto a screen, for instance (Hirstein, 2004).

4. There is no Cartesian theater. Is there a special place in the brain, where our conscious representations are displayed, as the self watches, swayed by desires and nudged by motives? Something like a theater perhaps, where what we see appears on a screen, and what we hear comes in via speakers? No, there are multiple, parallel processing streams, not a single, essential "stream of consciousness," and none of these processing streams contain anything like such a display area, according to this popular view. We are aware of only a tiny fraction of these streams, and that awareness is not like being in a theater. Dennett, who coined the term "Cartesian theater" (1991, p.17), has been the most constant critic of the idea behind it. "The idea of a special center in the brain," he says, "is the most tenacious bad idea bedeviling our attempts to think about consciousness" (1991, p.108).

Contrary to this, I will argue that the brain does contain something like a Cartesian theater. The brain's architectural plan includes processes that embody the highest-level representations. Executive processes have access to these processes, and this allows them to perform their operations on them. The late Francis Crick and his collaborator Christof Koch called the process in which information from the different sensory modalities is bound into a unified multimodal representation, a process of producing an "executive summary":

> The biological usefulness of visual awareness. . .is to produce the best current interpretation of the visual scene, in the light of past experience. . .and to make it available, for a sufficient time, to the parts of the brain that contemplate, plan and execute voluntary motor output.
>
> (Crick and Koch, 1995, p.121)

> A single, compact representation of what is out there is presented, for a sufficient time, to the parts of the brain that can choose among different plans of action.
>
> (Koch, 2004, p.234)

> Furthermore, once information is consciously accessible, it can be used to veto and suppress zombie [automatic] behaviors, actions, or memories inappropriate to the situation at hand.
>
> (Koch, 2004, p.234)

Baars (2005) explicitly employs the theatre metaphor, arguing that the contents of consciousness are "broadcast" to frontal processes that employ them to produce behavior.

The idea of a Cartesian theater runs contrary to the oft-repeated claim that there is not a single "place" where the highest-level mental activity takes place, but rather several parallel streams (e.g., in Dennett, 1991). Let me quickly list four considerations in favor of the substantial unity of the mind and its stream of consciousness: First, the mind's primary outputs are the body's actions. We only have one body, and the vast majority of

the time we are performing a single action, we are making coffee, picking up a book, washing our hands, and so on. Almost all of the body's muscles must work in parallel to achieve these actions, so at lower levels, the brain is doing many things. But there is nevertheless a high level at which one task is being accomplished. Second, speaking is a special type of action, and consists of the production of a single, serial stream of words. There needs to be a set of brain processes that can put our thoughts and mental states into a stream of words. Third, while there are several multimodal areas in the brain, they are all interconnected and these connections allow the brain to bind and unify these areas in a single conscious state. Fourth, the brain frequently uses a winner-take-all (and suppress losers) strategy in its computational operations (Koch, 2004, p.161). There is evidence that the brain suppresses memories called up but not strictly relevant to the ongoing task, for instance. Again, the simplest explanation of this is that the brain is working to produce a single, unified, consistent set of representations.

5. There are no special high-level control processes in the brain, i.e., no self. According to our everyday ways of thinking and talking, the mind contains more than the conscious thoughts, images, and other representations of the things around us. It contains more than the emotions, moods, and desires we experience. There is something else there, something that interacts with the representations, is influenced by the emotions, and moved by the desires. This something else apprehends our conscious perceptual states, elicits memories, meshes them with perceptions, and, taking account of emotions and motives, somehow distills this entire concoction into plans, intentions, and actions. We say, for instance, "I see it in my mind's eye," or "I analyzed my thought," or "I realized that I like the taste of pistachios," or "I can feel pain," and this seems to indicate that there is something in consciousness other than the image, the thought, the taste, and the pain. Traditionally, this something else has been called *the self.*

Philosophers who believe in privacy also tend to be skeptics about the existence of an inner, mental or psychological self. As we will see in Chapter 7, enlightenment philosopher David Hume's claim that he could not detect a self anywhere in his mind created a challenge to explain this that most philosophers believe has never been met. "Hume's denial," says Shoemaker (1996, p.3) "has been repeated by philosophers as different as Kant and Wittgenstein and has commanded the assent of the majority of subsequent philosophers who have addressed the issue." Most philosophers believe that anyone who believes that there is a self has probably committed what they call the *homunculus fallacy.* Someone commits this fallacy when, in the process of attempting an explanation of the mind, he invokes a process that has all of the mental abilities of a human being, and this is what the self is, they believe. "Homunculus" is Latin for "little man"; the self is nothing more than a little man hidden in the mind, the objection goes. Sigmund Freud, for instance, has been accused of committing this fallacy, since he sometimes describes his id, ego, and superego as if they were little people inside the mind, vying for control of it and the body. The self doesn't apprehend and react to our conscious states, according to the popular view, we, the full person apprehend and react to our conscious states. To remove the appearance of a self from remarks such as, "I am in pain," Ludwig Wittgenstein—perhaps

the most influential philosopher of the first half of the 20th century, and a powerful advocate of many of these popular views—suggested that we simply say, "There is pain" (1955). Lichtenberg (1990, p.189) likewise suggests that instead of saying "I am thinking," we should say, "It thinks," just as we say, "It is raining." "Are there entities," Dennett asks, "either *in* our brains, or *over and above* our brains, that control our bodies, think our thoughts, and make our decisions? Of course not! Such an idea is either empirical idiocy (James's 'pontifical neuron') or metaphysical claptrap (Ryle's 'ghost in the machine')" (1991, 413). "There is," asserts Dennett, "no Oval office in the brain" (1991, p.429).

Contrary to this, I will argue that the brain contains self-like processes. While there is no single brain area or process that achieves everything ascribed to the self, the set of executive processes function as a unified ensemble to disambiguate perception, correct memories, make decisions, plan actions, and many other mental processes. We describe the workings of these executive processes using the word "I," and their operations are in accord with the commonsense notion of the self as an active presence in the conscious mind. In Chapters 6, 7, and 8, I will develop the view that the brain's set of executive processes functions as a self, both in that it fits quite nicely what has historically been said about the self, and that the processes constitute a unified entity.

6. Introspection is not analogous to perception. It is not possible to divide the act of introspection into an object of perception and a perceiver. One reason why this is true is that there is no perceiver in the brain, i.e., no self, no homunculus. Searle affirms that:

> Where conscious subjectivity is concerned, there is no distinction between the observer and the thing observed, between the perception and the object perceived. The model of vision works on the presupposition that there is a distinction between the thing seen and the seeing of it. But for "introspection" there is simply no way to make this separation. Any introspection I have of my own conscious state is itself that conscious state.
>
> (Searle, 1992, p.97, see also 1992, pp.143–4)

Similarly, theories of introspection which are modeled on perception commit "a mistake of a fundamental sort," according to Heil (2003, p.237). Two different criticisms have been given of the idea that introspection either is a type of perception, or is at least closely analogous to it. In introspection, there is no separate subject, as Searle objects above. Second, perception is, by definition, of something external, otherwise we call it hallucination, imagery, or something like that.

Contrary to this, I will argue that introspection is strongly analogous to perception. Introspection involves something like a perceiver, and something like an object of perception. What we think of as introspection occurs when certain executive processes participate in certain types of causal relations with conscious representations residing in posterior cortical areas. The executive processes are analogous to the perceiver, and the representations are analogous to the objects of perception. Armstrong (1984) and Lycan (2006), also argue that introspection is indeed a species of perception. I will argue in Chapter 10 that each side is roughly half right. Searle is focusing on what I call the objective senses of our perception verbs, whereas Armstrong and Lycan are focusing on what I will call the notional senses of those verbs. I think Armstrong and Lycan are basically

correct, but my account of how this perception works and what it involves will be different from theirs.

7. There are no mental representations. Since representations, by definition, must be representations for someone, or to someone, to posit internal representations is to posit something or someone in the mind who the representations are for, i.e., a homunculus. It is wrong to conceive of representations as presentations to a homunculus, images projected on a screen in a Cartesian theater. Rather, the only representations are external objects, states, and events, such as photos, books, movies, and spoken sentences, according to this view.

Against this, I will argue that there are internal representations. Once it is clear that the executive processes are playing the role of the consumer of representations, we have a coherent reason for speaking of mental representations, and internal representation can be understood as analogous to external representation. In both cases, there is an object represented, a representation, and something (or someone) that uses the representation. I will argue, primarily in Chapter 12, that a simple and coherent account can be developed according to which there are not only representations, but mental representations. I accept the claim of the homunculus skeptics that representations need to be for somebody, or at least something, arguing in Chapter 5 that mental representations are for the brain's executive processes.

8. Colors are not internal, mental properties. There is a fascinating debate underway in both philosophy and neuroscience about where exactly the colors we experience exist. Are colors out there in the world, where they appear to be, or are they generated by our brains (or some combination of the two)? The issue is forced by the fact that the orthodox approach, according to which the color of an object is completely determined by the wavelengths of light reflected off of it (or the object's "spectral reflectance profile"), is subject to a several stubborn counterexamples. Benham's disk, for instance, is a small spinning top, colored half-black and half-white with black arcs in the white zone. Hold it in your hands and look at it and you won't see any colors. But put it on a table and give it a spin, and distinct circles of red, blue, green, and brownish-gold appear on the surface of the disk. The problem for the orthodox view is that none of the light coming off the disk has the frequencies that should correspond to those colors. Similarly, dreams that occur in color and colored mental images are also counterexamples to the orthodox theory, since they can be experienced in the absence of any light at all.

The existence of these and other counterexamples has led some scientists and philosophers to suggest that colors are not in fact possessed by the objects we see, or even by the light reflected off of them, but rather are created by the brain at a certain stage of visual processing. One reason, however, why many would balk at this sort of color internalism is that it again seems to presuppose a homunculus who sees the colors. After all, why should the brain go through the trouble to produce colors, since there is nothing in the brain to "perceive" them?

I will argue that colors are internal, mental properties, in Chapter 5. There are strong arguments against the idea that colors are properties of everyday objects. However, color

internalism is stymied by its inability to explain why the visual system would add color in the course of its processing. In the alternative view, the addition of color makes perfect sense. Colors are added in order to make stimuli salient to executive processes so that actions can be more effectively directed out into the world. The addition of color to the visual processing stream is an example of an engineering solution to the problems of life that the human brain evolved.

9. There is no filling-in. Each of your eyes has a rather large, roughly round, blind spot in the upper, outer part of its field of view. This occurs because the optic nerve exits the back of the eyeball at the place on the retina corresponding to the blind spot, and no light-sensitive cells exist there. Yet we are not aware of two large black spots. Why? Most scientists believe that this is because visual processes "fill in" this part of the conscious visual field using information from the areas surrounding the blind spot. But, according to this popular view, it is wrong to speak this way, because this gives the impression of something painting over the internal representation projected on the screen in the Cartesian theater. Dennett again: "The brain doesn't actually have to go through the trouble of 'filling in' anything. . .for no one is looking" (1991, p.127).

In my alternative view, conscious representations are filled in, for a specific reason. Filling-in is one of a number of processes that prepare conscious mental representations to interact with executive processes. In Chapter 5 I described a hypothesis, according to which the human brain evolved a cognitive system consisting of executive processes, limited in scope and function, which causally interact with highly processed representations. The unexplained phenomenon, filling-in of the blind spot, can be explained on this hypothesis: It was easier for the forces of evolution to fill in the visual field than to modify the executive processes that interact with it.

According to the cluster of popular views, the problem of the metaphysical nature of color, and the question of whether there is filling-in, are seen as two completely unrelated riddles. In the alternative framework, these problems have a common root: a failure to understand the role of executive processes in the mind/brain, as well as their need to have representations prepared for them. The three claims—no mental representations, no internal colors, and no filling-in—are motivated in many cases by the rejection of a self. They do not provide independent evidence for dismissing a notion of self, but they are consistent with it.

10. Folk psychology is largely false. "Folk psychology" is the name given to our everyday ways of talking about our mental lives, in which we use terms such as "believe," "see," "know," "remember," "love," and "imagine." Most of today's philosophers, with Searle as perhaps the main exception, hold some type of deflationary position about folk psychology, according to which it is either largely false (e.g., Churchland, 1979) or at least contains terms that are useful, but in the end fail to refer to anything real in the brain (e.g., Dennett, 1991). Just as all of our other everyday ways of talking about the world, folk physics, folk astronomy, folk chemistry, and so on, have proven to be fundamentally mistaken, folk psychology will also be shown to be filled with errors. The apparent fact that folk psychology allows us to describe our minds and explain our behavior does not

show that it is correct. Paul and Patricia Churchland are the primary spokespeople for this view. According to Paul Churchland:

> Our commonsense psychological framework is a false and radically misleading conception of the causes of human behavior and the nature of cognitive activity. On this view, folk psychology is not just an incomplete representation of our inner natures; it is an outright misrepresentation of our internal states and activities.
>
> (Churchland, 1984, p.43)

There are scientific experiments that appear to show that our sense of intentionally performing an action occurs *after* the action has already begun, so that the conscious intention only appears to be a cause of the action (Libet, 1993). What we are aware of is a mere *epiphenomenon*, like the puffs of smoke coming out of a steam locomotive; they are interesting to see, but they are not doing any of the work of moving the train. "The mind creates this continuous illusion" says Wegner, "it really doesn't know what causes its own actions" (2002, p.28). Gazzaniga believes that this sort of thing describes "a large part of our conscious reality":

> Behaviors are being continually exhibited, the origins of which may come from coherent, independent mental subsystems, and these actions have to be and are immediately interpreted by the verbal system. As a consequence, a large part of our sense of conscious reality, we believe, comes from the verbal system attributing cause to behavior.
>
> (Gazzaniga et al. 1977, p.1147)

Many authors believe that reports based on introspection are, for the most part, confabulations. The ordinary belief in introspection as a reliable method for discerning what goes on in our minds cannot be saved, according to the popular views on the matter. According to Dennett:

> . . .there are circumstances in which people are just wrong about what they are doing and how they are doing it. It is not that they lie in the experimental situation but that they confabulate; they fill in gaps, guess, speculate, mistake theorizing for observing. . . . They don't have any way of "seeing" (with an inner eye, presumably) the processes that govern their assertions, but that doesn't stop them from having heartfelt opinions to express.
>
> (Dennett, 1991, p.94)

In the alternative framework, I will develop throughout this book, folk psychology will be seen to be surprisingly accurate in the ways it describes mental states and events. In their eagerness to correct our everyday view, many philosophers have failed to attach sufficient weight to the extent, success, and longevity of folk psychology. In the alternative view, folk psychology contains a theoretical structure that is in general correct, and correct in many of its details. Folk psychology depicts conscious mental states as interactions between representations and executive structures, I will argue in Chapter 10. Our folk-psychological sense of self involves something in the mind, a self, interacting with mental representations to produce behavior. While the idea that there is a single entity or brain process that accomplishes all of these things may be a confabulation in response to social pressures, there nevertheless is a set of executive processes that accomplish all these things. Keeping folk psychology also saves our existing notions of intentional action and

responsibility for what we do, which many recent writers have argued will be seen to be fictions once we understand the brain (see, e.g., Gazzaniga, 1998; Wegner and Wheatley, 1999; Wegner, 2002).

The best way to be clear on these difficult questions about the mind is to be consistent with our existing concept of, and beliefs about, the mind. The question that gives rise to the problem of consciousness is a question about how the mind and its consciousness, *as we know them*, can be explained using only the language of the physical. We are experts in the use of the intricate web of concepts that make up folk psychology, such as *believe*, *think*, *remember*, *know*, *imagine*, and *love*. While our introspective reports can be mistaken, and the patterns in these mistakes are important and worth studying, an overemphasis on these errors has led several writers to argue that introspective reports are always, or routinely, false, or worse, that they always fail to refer to anything real. Contrary to this, in the alternative view, the forms of language we use to make introspective reports are tailored to match our brain structure. Introspective reports detail the causal relations between perceptual representations and executive processes. The structure of folk-psychological reports mirrors the basic functional structure of the brain. The sentence, "I am aware of my mental image," for instance, has a noun phrase corresponding to an executive process, and a second noun phrase corresponding to my conscious image, while the verb "aware" describes the kind of causal relation that holds between the two.

Conclusion

The ten popular views in the previous section support one another. We cannot separate our knowledge of our conscious states from their existence, because a perceptual theory of introspection is false, because there is no homunculus. And, there is no filling-in, because there is no self to fill in for. We report that we perceive a filled-in visual field, but this is because we are bad at introspection. The denial of the Cartesian theater, of the self, and of filling-in are inferentially connected in Dennett's thinking: "For whose benefit is all this animated cartooning [i.e., filling-in] being executed? For the audience in the Cartesian Theater. But since there is no such theater, there is no such audience" (1991, p.128). Perhaps the most widely held elements of these ten views are those espousing self-skepticism. Searle, for instance, is a self-skeptic, and adheres to the inseparability of the epistemology and metaphysics of conscious states, but is a realist about folk psychology and its notion of intentional action. The Churchlands abstain on the issue of whether the epistemology of conscious states can be separated from their ontology, but tend toward self-skepticism, and are, of course, famous for their skepticism about folk psychology, a view they call "eliminativism." Dennett espouses a variety of eliminativism and is also a self-skeptic (in the way I have described the self here), but does not adhere to the thesis of inseparability or, as we saw, of privacy. It is hard to fault people for taking the consistency of different subsets of these popular views as a sign that the entire subset is correct. Claims at this level are so difficult to verify by themselves, often the best one can do is build a consistent system of them, and hope that one of its tentacles meets with some concrete

counterexample. Barring this, one cannot help starting to think that the entire system is correct.

What I need to do next is provide an account of how one person might experience the mind of another in a way that shows that the set of ten alternative views broached earlier can form a coherent and useful framework for thinking about consciousness. In the next chapter, I will explain how I believe that the wall of privacy can be breached.