

PHILOSOPHY



Traditional and Experimental Readings

Fritz Allhoff
Ron Mallon
Shaun Nichols

New York • Oxford
OXFORD UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further Oxford University's objective of excellence in research, scholarship, and education.

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

Copyright © 2013 by Oxford University Press, Inc.

For titles covered by Section 112 of the US Higher Education Opportunity Act, please visit www.oup.com/us/he for the latest information about pricing and alternate formats.

Published by Oxford University Press, Inc.

198 Madison Avenue, New York, New York 10016

www.oup.com

Oxford is a registered trademark of Oxford University Press

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data

<to come>

Printing number: 9 8 7 6 5 4 3 2 1

Printed in the United States of America
on acid-free paper

Part II



Mind and Self



CHAPTER 4



Mental States

MARK PHELAN AND ERIC MANDELBAUM

INTRODUCTION

ONE MIGHT THINK THAT MAN IS a rational animal, as Aristotle (supposedly) did. Or one might prefer Oscar Wilde's formulation instead: "One is tempted to define man as a rational animal who always loses his temper when he is called upon to act in accordance with the dictates of reason" (Wilde 2007, 80). Either interpretation presupposes that people believe that things are a certain way, desire that they be some way or other, and sometimes intend to change the world to fit with their desires. To attribute *rationality* to people requires attributing *mentality* to people. It is these mental states, seeing that *p*, desiring that *q*, believing that *r*, intending that *x*, feeling that *y*, that are the focus of this section.

We will be specifically interested in three questions. The first is the question of whether mental states exist or not (call this the *ontological* question, for ontology is the study of what exists). There are multiple views one could hold: one could think that mental states exist just like chairs, tables, and electrons exist; one could think that mental states exist, but only in a loose sense—they exist in the same way that the average family exists (with its 1.86 children); or one could think that mental states just do not exist at all, like werewolves, unicorns, and honest politicians. Assuming that mental states exist, the second question concerns itself with what mental states are (call this the *identity* question, for it essentially inquires as to the identity of mental states). Assuming mental states exist at all, are they a wholly separate type of entity from the physical? Are mental states properties that naturally emerge from certain configurations of physical

states of the brain? Are mental states functional entities, ones that are what they are only in virtue of a certain role they play, like a mousetrap or a doorknob? The third is the question of how we know about the existence of mental states (call this the *epistemological* question, for epistemology is the study of knowledge). Do we know they exist because our best theories of human behavior—either philosophical, scientific, or commonsensical—require that they exist? Or do we know about their existence in some other way?

The importance of these questions stretches beyond the philosophy of mind; these are basic philosophical questions of how humans fit into the overall structure of the world. Is the mind, as physicalism holds, just made out of the same physical parts as other objects in the universe, like paper, rocks, scissors, and spiral nebulae? Or, as dualism suggests, is the mind such that it truly differs from all other extant substances? If the mind is special, then this implies that things with minds are, in a sense, special, too. Like the search for extraterrestrial intelligence, either answer we end up with will be a bit jarring but will also help us understand our place in the world a little better.

DESCARTES AND DUALISM

In his *Meditations on First Philosophy*, René Descartes argues in favor of the view that the mind is neither identical nor reducible to the body. Descartes' position on the mind is a variety of *substance dualism*; it holds that mental substances are a fundamentally different kind of thing than physical substances. Descartes arrives at this position through nonempirical, or *a priori*, reflection (reflection from the armchair, as it were). As he mentions at the beginning of *The Passions of the Soul* (p. XXX, this volume), Descartes embraces a methodological approach to philosophy according to which we start reasoning from what we cannot doubt. When it comes to considering the mind's relation to the physical body, this methodological approach recommends that we reason according to the following rule:

...anything we experience as being in us, and which we see can also exist in wholly inanimate bodies, must be attributed only to our body. On the other hand, anything in us which we cannot conceive in any way as capable of belonging to a body must be attributed to our soul. (p. XXX, this volume)

Descartes' argument for dualism relies on a connection between what we can conceive and what is possible. The argument adopts the principle that if we can conceive of X without conceiving of Y, then it must be the case that X is not identical to Y. Descartes goes on to argue that, since he can conceive of a body without a mind (e.g., a dead body) and a mind without a body (e.g., God, or an immortal disembodied soul), the body and mind must be distinct.

Whatever the theoretical strengths of substance dualism, Paul Bloom's short piece contends that it is an *intuitive* position. Bloom contends that we are natural-born dualists. In this and other work (e.g., Bloom 2004), Bloom argues for an innate, dedicated mental mechanism whose function is to detect others' mental states. The presence of such a mechanism is supposed to account for the ease with which people can

explain one another's behavior. Human behavior is much more complicated than the behavior, for example, of cars, yet almost everyone is adept at explaining the behavior of people by positing beliefs, desires, and other mental states, whereas explaining the behavior of cars is an acquired skill, one contained in a far smaller percentage of the population. What's more, children develop this ability to explain the behavior of others naturally, without any real instruction or even learning. Additionally, the small subgroup of people that have profound trouble explaining the behavior of others, people with autism, appear to have a developmental disorder. Based on this evidence, Bloom concludes that humans are mind readers from an early age; and he posits that when children learn about death, an intuitive dualism follows from this universal tendency to think in terms of mental states. Among other evidence in favor of natural-born dualism, Bloom refers to an experiment by Jesse Bering and David Bjorklund (2004). In these studies, preschool children are presented with a puppet show, in which Brown Mouse is having a terrible day. He's lost, sick, thirsty, and very sleepy. And things get even worse for Brown Mouse, because he's soon eaten by Mr. Alligator! After seeing the play, children are asked a series of questions. Though they tend to recognize that Brown Mouse won't ever need to go to the bathroom again and that his brain no longer works, they still think that he's hungry and that he wants to go home. In other words, kids tend to say that Brown Mouse's mental states persist, even when his biological processes cease. This is evidence that the belief in an immortal, disembodied mind is natural. As Bloom (2004) writes:

Young children do not know that they will one day die. But once they learn about the inevitable destruction of their body, the notion of an afterlife comes naturally. This is the most important consequence of seeing the world as Descartes did. (p. 208)

Regardless of its intuitive appeal and the philosophical arguments in favor of it, substance dualism is a startlingly bold hypothesis. It puts the mind outside of the physical world altogether. Descartes, however, did not intend to put the mind outside of the *causal* world altogether. Descartes thought the mind and the brain interact with one another. According to his view, neural activity leads to thought and feeling in our nonphysical minds; and the immateriality of a feeling is no bar on its influencing our behavior. Thus, Descartes' form of dualism is also a type of *interactionist dualism* (in addition to being a variety of substance dualism). The interactionist thinks that although the mind and body are separate, they can interact in causal ways. Since Descartes embraces the conjunction of substance dualism and interactionist dualism, we will refer to the conjunction of these two theses as *Cartesian dualism*.

In a lengthy correspondence with Descartes (several letters of which are reprinted in this volume), Princess Elisabeth of Bohemia raised an important challenge to Cartesian dualism. As she noted, it is quite difficult to see how the mind can interact with the body if the mind is necessarily a nonphysical thing. Causation seems to only hold between physical entities; in other words, causation only holds in the world of the physical. If the physical is a closed causal system, then how can a nonphysical entity, like the mind, interact with a physical entity? Causation seems to need some form of contact or, at least, some relationship over a distance (like gravitational pull), but how can a nonextended entity either make contact with or even stand at any distance

removed from a physical thing? As Cottingham et al. (1985) note, Descartes' reply to this question is as follows:

...that the body causes the soul to have feelings and passions, and the soul causes the body to move, through an inexplicable "union" between the soul and body—did not satisfy the princess. Nor was she satisfied when Descartes sought to answer her questions with vague moralizing and practical advice for the control of the passions. (325)

Descartes eventually obliged Elisabeth's request that he "define the passions, in order to know them better" (p. XXX, this volume). Relying on his previous groundbreaking anatomical research involving experimenting via vivisection to investigate muscular movement, among other empirical undertakings, Descartes wrote his final treatise, *The Passions of the Soul*, as a response to Elisabeth. Descartes surmised that the mind and the body must interact somewhere, and he concluded that the pineal gland was the most reasonable seat of the mind/body interactions. Anatomically, this was a reasonable guess given the limited understanding of neuroanatomy at the time, for the pineal gland is centrally located between the two hemispheres of the brain. Thus, it seems that the pineal gland might be the type of juncture that could affect the entire brain. However, this somewhat clever solution does not speak to the conceptual question that Elisabeth raised: How is it possible for a nonphysical entity to causally affect a physical one? If the physical world is a causally closed system, then how can a fundamentally nonphysical entity affect the physical world in any way?

Phelan et al. raise a further challenge for Cartesian dualism. One of the greatest promises of Cartesian dualism is purported to be the possibility that the nonphysical mind is protected from the ravages of the physical body. However, cases of brain damage indicate that physical injuries to the brain cause serious harm to the mind. Cartesian dualism needs to be able to explain why the nonphysical mind seems to be damaged by physical damage to the brain, for this seems to suggest that ravages to the physical body do indeed cause ravages to the mind.

Because of the interaction problem, the challenge of brain damage, and other problems, Cartesian dualism is rejected by most contemporary philosophers.

THE ELIMINATION OF THE MENTAL

If our intuitive theory of mind, Cartesian dualism, is fraught with so many conceptual problems, why not just abandon it, like other intuitive but conceptually problematic belief systems? The fact that an explanation comes naturally to us does not serve as justification for the explanation. For example, before receiving some rudimentary training in physics, people think that balls that are ejected from curved tubes will follow curved trajectories. This natural explanation of the path of the ball is part of our basic "folk physics." But such folk judgments are often the product of mental processes that rely on heuristics, or fast and imprecise rules, for dealing with a world in which numerous physical variables make timely and accurate physical predictions difficult to come by. The fact that folk explanations are overwhelmingly of a heuristic and ultimately false nature gives us some reason to reject them, or so claims the *eliminativist* Paul

Churchland. Like Bloom, Churchland claims that people employ an intuitive psychological theory—or *folk psychology*—that underlies their explanations of the behavior of others. However, Churchland goes on to claim that folk psychology is so badly mistaken that the very concepts it uses—belief, desire, intention—do not correspond to real entities in the world. Churchland likens folk psychology to failed physical and biological theories of yore like phlogiston theory and vitalism. Churchland argues that folk psychology is a stagnant theory, one that has not evolved in any meaningful way in eons. According to Churchland, folk psychology also fails to explain a wide array of data that should be in its explanatory purview. Churchland also charges folk psychology with being unable to integrate itself with what he sees as the emerging synthesis among various previously distinct disciplines. As the prototypical eliminative materialist, Churchland holds that not only is dualism false but also that beliefs, desires, and intentions do not exist.¹

ELIMINATIVISM'S RELIANCE ON ARGUMENTS FROM REFERENCE

What motivated Churchland to argue for the elimination of beliefs, desires, and intentions is the idea that folk psychological terms such as *belief* and *desire* don't actually refer to anything. If none of the terms used in a particular branch of science actually picks out any real things, then we have a good reason for getting rid of that science. Suppose, for example, that there were no such things as electrons, protons, or any of the fundamental posits of physics. In that case, we would have a pretty good reason for concluding that our science of physics was deeply mistaken and in need of rejection or fundamental reconceptualization. But, as Mallon et al. argue in their included piece, determining whether or not a scientific term actually refers to anything in the world is a nontrivial matter, and, importantly, it presupposes some theory of what it is for a scientific concept to refer to an object in the world.

Mallon et al. identify this argument for eliminativism as an *argument from reference*. In an argument from reference, a theorist first adopts a particular theory of how terms refer. Churchland and other eliminativists, for example, assume that concepts refer to whatever satisfies the description people typically associate with that term. Next, the theorist who is using an argument from reference goes on to show that, if the adopted theory of reference is correct, the terms of whatever scientific or philosophical theory they are discussing have some specific property. Churchland and other eliminativists, for instance, contend that if our terms refer descriptively, then the terms used in our folk psychological explanations fail to refer to anything, because nothing in the world satisfies the descriptions people associate with beliefs, desires, and so on. Finally, the proponent of an argument from reference goes on to draw some philosophically significant conclusion. For example, Churchland and other eliminativists conclude that folk psychological theories must be rejected because their terms (belief, desire, etc.) fail to map onto the actual world.

Arguments from reference, such as Churchland's argument for the elimination of mental states, only work if the particular theory of reference they invoke is the correct theory of reference. But how do we determine if a particular theory of reference is

correct? As Mallon et al. point out, “the correct theory of reference...is commonly thought to be constrained by our intuitions about...reference...in actual and fictional cases” (p. XXX, this volume). This method of settling on *the* correct theory of reference seems to assume that intuitions about actual and fictional cases will be more or less universally shared. However, as Mallon et al. conclude from a review of some recent experimental work, there is good reason to conclude that there may be cross-cultural and individual differences between people’s intuitions about important fictional cases. If so, our intuitions fail to show that one theory of reference is *the* correct one, and Churchland’s conclusion, among the conclusions that assume a particular theory of reference, fail to follow.

FODOR, FUNCTIONALISM, AND THE PERSISTENCE OF THE ATTITUDES

Mallon et al. try to undercut eliminativism by going negative: they argue that a crucial premise in the eliminativist argument, the premise relying on a theory of reference, is unsupportable. However, one can also try to stave off eliminativism by offering a positive account of what mental states are. *Functionalism* attempts to offer such an account.

Between such stark positions as dualism and eliminativism lies functionalism. The functionalist thinks that states such as believing, hoping, wanting, doubting, and so on, are functional states—that is, states that function in a certain way and are identified by their functions. For the functionalist, what makes something a belief is not that it has a certain structure (e.g., it does not matter that psychological states are composed out of neurological states); rather, what matters for the identity of a psychological state is that it serves a certain role in one’s psychological economy. The functionalist points out that psychological states are like lots of other things that we identify based on what they do. What makes something a mousetrap, for example, is that it serves a certain function, that is, trapping mice. Be it a snap trap or glue trap, it’s the role it plays in our homes that is important. What is not important to identifying a mousetrap is the specific materials out of which it is built; a mousetrap can be built out of wood, metal, plastic, and so on. Likewise, the functionalist holds that what makes something, for example, a belief, is that it serves a certain role in a mental economy.

Jerry Fodor attempts to secure the ontological footing of psychological states in just this way. For Fodor, as things stand in our world, psychological states happen to always be instantiated by neurological states; that is, every given instance of a psychological state is also an instance of a neurological state. However, Fodor thinks that different kinds of psychological states, such as beliefs, are not just identical to different kinds of neurological states. To put it crudely, there isn’t a belief neuron, nor is believing always achieved by some particular part of the brain. Instead, a mental state like “believes” is realized by lots of different neurological states (and could even, in principle, be realized by nonneurological [e.g., silicon] robot CPU states). Again, what is ultimately important is the role a state plays, not what it is made of.

On the picture just sketched, the functionalist arrives at a middle position between the extremities of embracing a dualism and eliminating psychological states altogether.

Since most functionalists think that the things that happen to play the psychological state role are all, as a matter of fact, neurological states, most functionalists don't need to posit any nonphysical realm of the mind, as Descartes did. On the other hand, because psychological states are not just neurological states, one can still talk about them in terms of their distinctively psychological roles.

Of course, one might still ask the functionalist, why posit psychological states at all? It is this question that Fodor's paper takes up. Fodor is a full-fledged psychological realist: he thinks psychological states are mental representations that have real causal powers, just as tables, chairs, and protons do. Taking beliefs, desires, and the like to have real causal powers allows us to explain vast amounts of human behavior. Fodor argues that our theory of mind works so well that it is easy to overlook its innumerable successes. Just by uttering certain words, we can predict disparate people's behavior despite numerous individual differences (and these predictions can be quite complicated, traversing a vast amount of time and space). Fodor takes such predictive accuracy as a strong reason for believing in the existence of mental states. After all, how could we account for such predictive accuracy if mental states, the entities we use in such predictions, were actually nonexistent? And note that, in other areas of science when we use a certain entity for prediction and explanation, we assume that entity really exists. When we speak of gravitational fields causing smaller celestial bodies to be attracted to larger celestial bodies, we take it that such a gravitational field actually exists. Part of our reason for believing in its existence is that by positing such a phenomenon, we can explain our observations. Our understanding and predicting of behavior involves positing mental states, so Fodor thinks we have good reason to allow such states into our ontology, as we would in any other investigative inquiry.

REAL PATTERNS AND INTERPRETIVE STANCES

Just as the functionalist takes a more moderate ontological stance than the dualist and eliminative materialist, Dennett seeks to take an even more concessive position. Dennett is an *interpretationist*. As opposed to the functional realism of someone like Fodor, Dennett proposes that mental states are not necessarily inner, causal states of people. Whereas the functionalist thinks that beliefs and desires give us a mechanistic explanation of behavior, the interpretationist only goes so far as to claim that intentional states are tools we use in predicting behavior. For the interpretationist, beliefs and desires are like centers of gravity and the average American family. The average American family needn't exist (and in fact couldn't exist if it has only 1.86 children), but it could still be quite useful to use in predicting certain facts. For example, were someone to offer you \$100 to predict how many children are in the randomly chosen Jones family, it would be reasonable to use your knowledge of the average American family, as well as your knowledge that there are no 86% children, to generate a guess of 2. Although such a concept is useful, that does not imply that it refers to an existing entity.

Dennett describes explaining people's behavior in terms of *intentional states* (another term for mental states such as beliefs and desires) as taking the "intentional stance" (Dennett, 1987) toward a person (or object, for that matter). Taking the

intentional stance involves predicting behavior by positing beliefs and desires that it is rational for the person (or object) in question to have. For Dennett, there is no further fact of the matter about what beliefs the person actually has; rather, we assign beliefs against an assumption of rationality. Hence, the interpretationist does not take belief ascription to be a merely descriptive endeavor like most scientific observations are. For a typical example of scientific observation, consider how in comparative psychology experimenters often use experimental paradigms where they observe how long, for example, a rhesus macaque monkey looks directly at a stimulus. Coding how long the monkeys looked in a certain direction is, when all goes well, merely a descriptive endeavor. Since, for Dennett, belief ascription does not follow this strictly descriptive route, belief ascription does not seem properly scientific. Dennett concludes that, contra Churchland, beliefs exist but not necessarily in the “industrial strength” way of Fodor. Instead, Dennett thinks that they describe certain patterns of behavior. Beliefs do not necessarily carve the world at its joints, but they are helpful in picking real patterns that emerge from neural activity. In sum, Dennett is a materialist who wants to stave off eliminativism and Fodor’s industrial-strength realism by embracing beliefs as an abstract description of a pattern of underlying neural activity.

FOLK PSYCHOLOGY AS A FULL-BLOODED THEORY

One recurring theme of this discussion has been that people use their ascriptions of mental states to predict other people’s behavior to some greater or lesser extent. But there are other questions we might ask about the process of mental state ascription. For instance, many contemporary philosophers and cognitive scientists have asked what the specific psychological mechanisms are that allow us to attribute mental states to others. *Theory theory* provides one answer to this question. It is the position that people understand other people’s minds on the basis of an unconscious, tacit theory about how people’s minds work. In particular, the theory theory of Gopnik and Wellman claims that children gain an understanding of others’ minds by using a tacit theory. The theory theory theorist sees the child as a little scientist, gathering evidence, testing hypotheses, and revising his theories of the mind accordingly.

Gopnik and Wellman’s argument in favor of theory theory has two main moves. First, they analyze what it takes for something to count as a theory, and then they marshal evidence from developmental psychology to support their contentions. They claim that theories posit abstract entities to interpret, predict, and explain certain phenomena. These entities should cohere in a lawful way with one another and with the observed phenomena. Theories aren’t just static bodies of entities in lawful correlation with one another; instead, theories are dynamic entities, entities that change as new evidence comes in. For example, when a theory’s hypotheses are disproven, they are replaced by better, more confirmed hypotheses.

After delineating some characteristics of theories, Gopnik and Wellman delve into the developmental data to show that the child’s theory of mind undergoes certain transitions. For example, Gopnik and Wellman claim that the two-year-old only posits desires and perceptions. However, the three-year-old countenances a greater number

of mental states such as pretending, knowing, and dreaming. Yet the three-year-old still has trouble with belief, particularly false belief. The three-year-old does not have a robust understanding of how someone can misrepresent some state of affairs. By the time the three-year-old matures to a five-year-old, his or her theories too have matured. The five-year-old understands representations and misrepresentations and can, in particular, make sense of false belief. In this way, Gopnik and Wellman interpret the child as holding a theory that gains theoretical depth and explanatory power as the child develops. Between two and five years old, the child's theory gains more abstract entities and the child understands the relations between such entities and the relations between the entities and the world better as time goes on.

THE PERSON AS SCIENTIST, THE PERSON AS MORALIST

Joshua Knobe also addresses the topic of how we use mental states to explain one another's behavior. Knobe's paper takes aim at a Gopnik and Wellman-style theory theory. Gopnik and Wellman suggest that the way children learn about others' minds is akin to the way that scientists develop theoretical frameworks. In opposition, Knobe contends that people's everyday reasoning is unlike scientific reasoning in that everyday reasoning is suffused with moral considerations. Contra Gopnik and Wellman, Knobe has developed the view that "there is literally a non-conscious process whereby moral judgments influence intuitions about intentional action which then serve as input to the mechanisms underlying intuitions about [other people's mental states]" (2007, 98). (In the included selection, Knobe reviews some of the evidence supporting this claim.) This has the result that the attribution and use of mental states for the purposes of explaining action itself rests in part on moral considerations. Knobe concludes that when we explain one another's behaviors, we are not merely tough-minded scientists, as Gopnik and Wellman's theory suggests. When we want to account for our attributions of mental states, we must also recognize that the theory that drives those attributions is driven in part by moral considerations.

Note

1. We focus on beliefs, desires, and intentions—or *propositional attitudes*—as opposed to mental states in general, because Churchland believes that sensations needn't necessarily be eliminated. Churchland thinks sensations will be identified with brain states and ultimately reduced to them, whereas he thinks we need to eliminate all the propositional attitudes from our ontology.

References

- Bloom, P. *Descartes Baby: How the Science of Child Development Explains What Makes Us Human*. New York: Basic Books, 2004.
- Cottingham, J., R. Stoothoff, and D. Murdoch. *The Philosophical Writings of Descartes*. Cambridge: Cambridge University Press, 1985.
- Dennett, D. "True Believers." In *The Intentional Stance*. Cambridge, MA: MIT Press, 1987.
- Descartes, R. *The Passions of the Soul* I XXXI. In *The Philosophical Works of Descartes*. Vol. 1, trans. and ed. E. Haldane and G. Ross. Cambridge: Cambridge University Press, 1969.

Knobe, J. "Reason Explanation in Folk Psychology." *Midwest Studies in Philosophy* 31, no. 1 (2007): 90–106.

Wilde, O. *The Critic as Artist (Upon the Importance of Doing Nothing and Discussing Everything)*. New York: Mondial, 2007.

Suggestions for Further Reading

Alexander, J., R. Mallon, and J. M. Weinberg. "Accentuate the Negative." *Review of Philosophy and Psychology* 1, no. 2 (2010): 297–314.

Aydede, M., and G. Guzeldere. "Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness." *Nous* 39, no. 2 (2005): 197–255.

Baron-Cohen, S. *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press/Bradford Books, 1995.

Bealer, G. "Self-Consciousness." *Philosophical Review* 106, no. 1 (1997): 69–117.

Bering, J., and D. Bjorklund. "The Natural Emergence of Afterlife Reasoning as a Developmental Regularity." *Developmental Psychology* 40 (2004): 217–233.

Chalmers, D. *The Conscious Mind*. Oxford: Oxford University Press, 1996.

Chalmers, David J. "Phenomenal Concepts and the Knowledge Argument." In *There's Something About Mary: Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument*, ed. P. Ludlow, Y. Nagasawa, and D. Stoljar. Cambridge, MA: MIT Press, 2004.

Dennett, D. "The Unimagined Preposterousness of Zombies." *Journal of Consciousness Studies* 2, no. 4 (1995): 322–326.

Dennett, Daniel C. "What Robomary Knows." In *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, ed. T. Alter and S. Walter. New York: Oxford University Press, 2006.

Descartes, R. *Meditations on First Philosophy*. Englewood Cliffs, NJ: Prentice Hall, 1960.

Fodor, J. "Special Sciences (or: The Disunity of Science as a Working Hypothesis)." *Synthese* 28 (1974): 77–115.

Jackson, F. "Epiphenomenal Qualia." *Philosophical Quarterly* 32 (1982): 127–136.

Jackson, F. "What Mary Didn't Know." *Journal of Philosophy* 83 (1986): 291–295.

Knobe, J., and J. Prinz. "Intuitions about Consciousness: Experimental Studies." *Phenomenology and the Cognitive Sciences* 7, no. 1 (2008): 67–83.

Knobe, J., and S. Nichols, eds. *Experimental Philosophy*. New York: Oxford University Press, 2008.

Kripke, S. *Naming and Necessity*. Oxford: Blackwell, 1980.

Levine, J. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64 (1983): 354–361.

Lewis, D. "What Experience Teaches." In *Mind and Cognition*, ed. W. Lycan. Oxford: Wiley-Blackwell, 1990.

Lycan, W. G. "Kripke and the Materialists." *Journal of Philosophy* 71 (1974): 677–689.

Lycan, W. G. "Perspectival Representation and the Knowledge Argument." In *Consciousness: New Philosophical Perspectives*, ed. Q. Smith and A. Jokic. New York: Oxford University Press, 2003.

Machery, E., R. Mallon, S. Nichols, and S. P. Stich. "Semantics, Cross-cultural Style." *Cognition* 92, no. 3 (2004): B1–B12.

Nichols, S., and S. Stich. *Mindreading: An Integrated Account of Pretence, Self-Awareness and Understanding Other Minds*. Oxford: Oxford University Press, 2003.

Prinz, J. "Level-Headed Mysterianism and Artificial Experience." *Journal of Consciousness Studies* 10, no. 4–5 (2003): 111–132.