



PHILOSOPHY OF  
**MIND**

---

Classical and  
Contemporary Readings

David J. Chalmers

New York Oxford  
OXFORD UNIVERSITY PRESS  
2002

PHILOSOPHY OF  
**MIND**

---

Classical and  
Contemporary Readings

David J. Chalmers

New York    Oxford  
OXFORD UNIVERSITY PRESS  
2002

Oxford University Press

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai  
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata  
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi  
São Paulo Shanghai Singapore Taipei Tokyo Toronto

and an associated company in Berlin

Copyright © 2002 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.  
198 Madison Avenue, New York, New York, 10016  
<http://www.oup-usa.org>

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**

Chalmers, David John, 1966–

Philosophy of mind : classical and contemporary readings / David J. Chalmers.

p. cm.

Includes bibliographical references.

ISBN 0-19-514580-1 (hardback : alk. paper)—ISBN 0-19-514581-X (pbk. : alk. paper)

1. Philosophy of mind. I. Title.

BD418.3 .C435 2002

128'.2—dc21

2002072403

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

*Printed in the United States of America  
on acid-free papers*

# PREFACE

What is the mind? What is the relationship between mind and body? Is the mind the same as the brain? How can the mind affect the physical world? What is consciousness? Could a purely physical system be conscious? Can we explain subjective experience in objective terms? How does the mind represent the world? What is the nature of belief and desire? What is the relationship between consciousness and representation? Is the mind in the head or in the environment? What can we know about other minds, in humans, animals, and machines? What is the self?

These are some of the central questions in the philosophy of mind. This book is a collection of articles addressing them. If the book has a thematic focus, it is on the many aspects of the mind-body problem: What is the relationship between mind, brain, and body and between the mental and the physical? This is unquestionably the central problem in the philosophy of mind, and it ramifies into any number of different questions, concerning different aspects of this relationship and concerning different aspects of the mind. The articles in this book address these questions from many different angles.

This collection has three main parts, representing what are arguably the three main streams in the philosophy of mind. The first concerns foundational questions about the metaphysics of mind: What is the nature of the mind, and what is the relationship between the mental and the physical? The second concerns questions about consciousness: What is the place of consciousness in the natural world, is consciousness a physical process, and how can consciousness be understood? The third concerns questions about mental content: How can the mind represent the world, what is the nature of thought, and are the contents of our thoughts determined by the brain or by the environment? In addition, there is a brief fourth part addressing the problems of other minds, personal identity, free will, and artificial intelligence.

The collection includes both classical articles that make up much of the standard history of the field and contemporary articles that represent recent directions in the area. Much of the classical background to recent debates, from Descartes' dualism to various twentieth-century forms of materialism, can be found in the first part of the book. The second and third parts concentrate largely on material from the last few decades, with a good representation of material at the edge of current research. The book contains a combination of highly accessible articles and more sophisticated articles, so it should be suitable for use in undergraduate courses at all levels as well as in graduate courses. I hope that the book will also be interesting to general readers interested in these issues.

The philosophy of mind has become an enormous and diverse area of research in recent years, and it is impossible to cover the entire field in a book like this. In general, I have aimed for depth of coverage in the central areas of the field, but difficult omissions have had to be made. In particular, one cannot do justice to the thriving and sprawling area of the philosophy of cognitive science as a mere section in a book like this. Instead, this topic will be covered in a forthcoming companion volume, *Philosophy of Cognitive Science*.

I have written introductions to each of the parts, giving relevant background for the material in those parts as well as pointers for further reading. I have also put together a Web site for this book containing links to relevant online material, including extensive bibliographies and online articles and reference works. This Web site will be continually updated to cover recent developments in the field. Readers are encouraged to consult it at <http://consc.net/pom.html>.

I would like to thank Robert Miller of Oxford University Press for inviting me to put this book together and for all his help on the editorial front. Brad Thompson was a great help in

chasing down permissions and in preparing the manuscript. Thanks are due also to Fiona Cowie, George Graham, Jaegwon Kim, and two other reviewers for their helpful comments on the contents and organization of the book. Finally, I owe a debt to the editors of previous anthologies in the area—especially Ned Block (*Readings in the Philosophy of Psychology*), David Rosenthal (*The Nature of Mind*), and William Lycan (*Mind and Cognition*)—for their example.

# CONTENTS

Preface	xi
---------	----

---

## 1 Foundations 1

---

### A. Dualism

1 <b>Meditations on First Philosophy (II and VI)</b>	10
René Descartes	
2 <b>Passions of the Soul (Excerpt)</b>	21
René Descartes	
3 <b>On the Hypothesis That Animals Are Automata, and Its History (Excerpt)</b>	24
Thomas H. Huxley	
4 <b>An Unfortunate Dualist</b>	31
Raymond M. Smullyan	

---

### B. Behaviorism

5 <b>Descartes' Myth</b>	32
Gilbert Ryle	
6 <b>Psychology in Physical Language (Excerpt)</b>	39
Rudolf Carnap	
7 <b>Brains and Behavior</b>	45
Hilary Putnam	

---

### C. The Identity Theory

8 <b>Is Consciousness a Brain Process?</b>	55
U. T. Place	
9 <b>Sensations and Brain Processes</b>	60
J. J. C. Smart	
10 <b>The "Mental" and the "Physical" (Excerpt)</b>	68
Herbert Feigl	

---

### D. Functionalism

11 <b>The Nature of Mental States</b>	73
Hilary Putnam	

12	<b>The Causal Theory of the Mind</b> D. M. Armstrong	80
13	<b>Psychophysical and Theoretical Identifications</b> David Lewis	88
14	<b>Troubles with Functionalism</b> (Excerpt) Ned Block	94
15	<b>Pseudonormal Vision: An Actual Case of Qualia Inversion?</b> Martine Nida-Rümelin	99
<hr/>		
<b>E. Other Psychophysical Relations</b>		
16	<b>Mechanism and Its Alternatives</b> (Excerpt) C. D. Broad	106
17	<b>Mental Events</b> Donald Davidson	116
18	<b>Special Sciences (or: The Disunity of Science as a Working Hypothesis)</b> Jerry A. Fodor	126
19	<b>Multiple Realization and the Metaphysics of Reduction</b> Jaegwon Kim	135
20	<b>From Supervenience to Superdupervenience: Meeting the Demands of a Material World</b> (Excerpt) Terence Horgan	150
21	<b>Finding the Mind in the Natural World</b> Frank Jackson	162
<hr/>		
<b>F. Mental Causation</b>		
22	<b>The Many Problems of Mental Causation</b> (Excerpt) Jaegwon Kim	170
23	<b>Mental Causation</b> Stephen Yablo	179
<hr/>		
<b>2 Consciousness</b>		197
<hr/>		
<b>A. General</b>		
24	<b>Concepts of Consciousness</b> Ned Block	206
25	<b>What Is It Like to Be a Bat?</b> Thomas Nagel	219
26	<b>Quining Qualia</b> Daniel C. Dennett	226
27	<b>Consciousness and Its Place in Nature</b> David J. Chalmers	247

---

**B. The Knowledge Argument**

- |    |   |     |
|----|---|-----|
| 28 | <b>Epiphenomenal Qualia</b>               | 273 |
|    | Frank Jackson                             |     |
| 29 | <b>What Experience Teaches</b>            | 281 |
|    | David Lewis                               |     |
| 30 | <b>Phenomenal States (Second Version)</b> | 295 |
|    | Brian Loar                                |     |
| 31 | <b>Two Conceptions of the Physical</b>    | 311 |
|    | Daniel Stoljar                            |     |
- 

**C. Modal Arguments**

- |    |   |     |
|----|---|-----|
| 32 | <b>Naming and Necessity (Excerpt)</b>   | 329 |
|    | Saul A. Kripke  |     |
| 33 | <b>Imaginability, Conceivability, Possibility and the Mind-Body Problem (Excerpt)</b> | 334 |
|    | Christopher S. Hill   |     |
| 34 | <b>Rigid Designators and Mind-Brain Identity (Excerpt)</b>                            | 341 |
|    | Grover Maxwell  |     |
- 

**D. The Explanatory Gap**

- |    |  |     |
|----|--|-----|
| 35 | <b>Materialism and Qualia: The Explanatory Gap</b>           | 354 |
|    | Joseph Levine  |     |
| 36 | <b>The Rediscovery of Light</b>                              | 362 |
|    | Paul M. Churchland   |     |
| 37 | <b>Conceptual Analysis, Dualism, and the Explanatory Gap</b> | 371 |
|    | Ned Block and Robert Stalnaker                               |     |
| 38 | <b>Can We Solve the Mind-Body Problem?</b>                   | 394 |
|    | Colin McGinn   |     |
- 

**E. Higher-Order Thought and Representationalism**

- |    |   |     |
|----|---|-----|
| 39 | <b>Explaining Consciousness</b>                               | 406 |
|    | David M. Rosenthal  |     |
| 40 | <b>Conscious Experience</b>                                   | 422 |
|    | Fred Dretske  |     |
| 41 | <b>Sensation and the Content of Experience: A Distinction</b> | 435 |
|    | Christopher Peacocke  |     |
| 42 | <b>Visual Qualia and Visual Content Revisited</b>             | 447 |
|    | Michael Tye   |     |
| 43 | <b>Introspection and Phenomenal Character</b>                 | 457 |
|    | Sydney Shoemaker  |     |



<b>3</b>	<b>Content</b>	473
<hr/>		
<b>A.</b>	<b>The Nature of Intentionality</b>	
44	<b>The Distinction between Mental and Physical Phenomena</b> (Excerpt) Franz Brentano	479
45	<b>"Intentional Inexistence"</b> (Excerpt) Roderick M. Chisholm	484
46	<b>A Recipe for Thought</b> Fred Dretske	491
✓ 47	<b>Biosemanantics</b> Ruth Garrett Millikan	500
48	<b>Reasoning and Representing</b> Robert Brandom	509
49	<b>The Intentionality of Phenomenology and the Phenomenology of Intentionality</b> Terence Horgan and John Tienson	520
<hr/>		
<b>B.</b>	<b>Propositional Attitudes</b>	
50	<b>Empiricism and the Philosophy of Mind</b> (Excerpt) Wilfrid Sellars	534
51	<b>Propositional Attitudes</b> Jerry A. Fodor	542
✓ 52	<b>True Believers: The Intentional Strategy and Why It Works</b> Daniel C. Dennett	556
53	<b>Eliminative Materialism and the Propositional Attitudes</b> Paul M. Churchland	568
<hr/>		
<b>C.</b>	<b>Internalism and Externalism</b>	
54	<b>The Meaning of "Meaning"</b> (Excerpt) Hilary Putnam	581
55	<b>Individualism and the Mental</b> (Excerpt) Tyler Burge	597
56	<b>The Components of Content</b> (Revised Version) David J. Chalmers	608
57	<b>Anti-Individualism and Privileged Access</b> Michael McKinsey	634
58	<b>What an Anti-Individualist Knows A Priori</b> Anthony Brueckner	639
59	<b>The Extended Mind</b> Andy Clark and David J. Chalmers	643

---

<b>4</b>	<b>Miscellaneous</b>	653
60	<b>Reductionism and Personal Identity</b> Derek Parfit	655
61	<b>Freedom and Necessity</b> A. J. Ayer	662
62	<b>Analogy</b> Bertrand Russell	667
63	<b>Can Computers Think?</b> John R. Searle	669

# 1

# Foundations

The articles in this part address foundational questions about the nature of the mind and about the relationship between the mental and the physical. Many of these questions concern the nature of mental states: states such as seeing red, feeling pain, experiencing anger, and desiring happiness. What is the nature of a mental state? And how are mental states related to physical states, such as states of the brain, of one's body and behavior, and of the physical world more generally?

Traditionally, views on these issues can be divided into two main classes. *Dualist* views hold that the mind is quite distinct from the body and the brain (although they may be associated in some fashion), and/or that mental states are fundamentally distinct from physical states. *Materialist* views hold that the mind is itself a broadly physical entity, and/or that mental states are derivative on physical states. There also exist *idealist* views, according to which physical states are derivative on mental states, but these will not be as central here. The papers in this part discuss many varieties of dualist and materialist views, as well as other foundational questions about relation between the mental and the physical.

## A. Dualism

Dualist views come in two main varieties. *Interactionism* holds that the mental and physical are fundamentally distinct but interact in both directions: Physical states affect mental states, and mental states affect physical states. *Epiphenomenalism* holds that the mental and physical are fundamentally distinct and that physical states affect mental states, but denies that mental states affect physical states.

In the history of philosophy, the most important dualist view is the interactionism of René Descartes. Descartes' most important work is his *Meditations on First Philosophy*. This is a series of six meditations, the second and sixth of which are reproduced here as chapter 1. In the first meditation, Descartes attempts to cast doubt on all of his beliefs and finds that he cannot be certain that the external world exists. In the second meditation, Descartes finds that there is one thing he can be certain of: his own mind, and thus his own existence ("I think, therefore I am"). He concludes that he is fundamentally "a thing that thinks." In the third through fifth meditations, Descartes infers the existence of God and uses this to justify his belief in the external world (since God would not deceive him). In the sixth meditation, Descartes reflects on the differences between the mental and the physical and concludes that they are fundamentally distinct. He uses a number of arguments here: One can be certain about the mental but not about the physical; the

mind is indivisible while any physical entity is divisible; and most famously, one can imagine oneself existing without a body, so one must be distinct from one's body and likewise from any physical entity.

The *Meditations* argue for the distinctness of mind and body but do not say much about their relationship. This question is addressed in more depth in Descartes' *Passions of the Soul* (chapter 2), which discusses the interaction between mind and body. Humans have a rational soul, which receives perceptions as "passions" from the brain and performs actions through acts of will that affect the brain. Signals are passed between the brain and the soul via the pineal gland (a small gland centrally located in the brain). On this picture, mind and body involve separate substances but interact in both directions.

Descartes' ideas about the mind have been highly influential, but today they are widely rejected. Many objections have been raised to the idea that a nonphysical mind controls the movement of a physical body. It is not clear just how a nonphysical substance and a physical substance can interact: The idea that the pineal gland mediates this interaction has long since been rejected on physiological grounds, and it is unclear whether any better causal nexus could exist. Further, it is often held that this interaction cannot be reconciled with physics, which postulates a closed network of physical interactions, with no room for a nonphysical mind to play any role.

In reaction to objections of this sort, some have embraced epiphenomenalism, retaining the distinctness of mind and body but denying any causal role for mind in the physical world. Such a view is put forward by Thomas Huxley (chapter 3). Huxley addresses Descartes' view that nonhuman animals are mere automata, whose behavior is controlled entirely by their brain and which lack minds altogether. Huxley suggests that contemporary evidence favors the first aspect of this view, but does not favor the second: That is, animals' behavior is controlled entirely by their brain, but they have minds nevertheless. On this view, the mind is a sort of by-product of the brain that has no effect on it. At the end of his article, Huxley suggests that the same goes for humans.

Epiphenomenalism has the advantage of being easier to reconcile with science than interactionist dualism, but it has the disadvantage of running strongly counter to common sense. Intuitively, it is hard to accept that our thoughts and feelings have no effect on our behavior. Another problem is raised by Raymond Smullyan in his short fable "An Unfortunate Dualist" (chapter 4): If mind has no effect on behavior, then it has no effect on what we say about the mind, so it seems that one could remove the mind and we would go on talking about it just the same. Smullyan raises the issue as a problem for dualism in general, but it is particularly pressing for an epiphenomenalist.

### FURTHER READING

Objections to interactionism and epiphenomenalism are discussed in more detail by Kim (chapter 22) and Chalmers (chapter 27). Chalmers gives a limited defense of both interactionism and epiphenomenalism, while Jackson (chapter 28) defends an epiphenomenalist view. Elsewhere, interactionist views are defended by Foster (1991), Hodgson (1991), Popper and Eccles (1977), and Swinburne (1986), while an epiphenomenalist view is defended by Robinson (1988) and to a limited extent by Chalmers (1996). A relevant collection is Smythies and Beloff (1989).

Chalmers, D. J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.

Foster, J. 1991. *The Immaterial Self: A Defense of the Cartesian Dualist Conception of Mind*. Routledge.

Hodgson, D. 1991. *The Mind Matters: Consciousness and Choice in a Quantum World*. Oxford University Press.

Popper, K., and Eccles, J. 1977. *The Self and Its Brain: An Argument for Interactionism*. Springer.

- Robinson, W. S. 1998. *Brains and People: An Essay on Mentality and Its Causal Conditions*. Temple University Press.
- Smythies, J. R., and Beloff, J. (eds.). 1989. *The Case for Dualism*. University of Virginia Press.
- Swinburne, R. 1986. *The Evolution of the Soul*. Oxford University Press.

## B. Behaviorism

In the second half of the twentieth century, dualism was widely rejected, and many different forms of materialism were explored. This was both a reaction to the problems of dualism and a product of the success of physical explanations in many different domains.

Gilbert Ryle's 1949 book *The Concept of Mind* is recognizably the antecedent of much recent work in the philosophy of mind. This book argues against dualist views and puts forward a positive view of its own. Included here is the first chapter of the book, "Descartes' Myth" (chapter 5). As the title suggests, this chapter is largely directed against the dualism of Descartes. He accuses Descartes and others of subscribing to the "dogma of the ghost in the machine" and suggests that these views rest on a "category mistake" in posing questions about the relationship between mind and body. The mind is not to be seen as something distinct from the body and steering it from the inside, but as an aspect of the body's own activities.

Ryle's positive views (developed in other chapters of his book) are subtle and hard to summarize, but one strand in these views seems to involve a sort of *behaviorism*: roughly, the view that the mind is an aspect of behavior. On this view, to be in a given mental state (such as pain) is to be in a certain behavioral state (such as wincing), or at least to have a disposition to behave in certain ways (such as the disposition to express pain if queried). Thus mind is seen as a public aspect of human activity, rather than as a private inner aspect.

This sort of behaviorism is more explicit in Rudolf Carnap's "Psychology in Physical Language" (chapter 6). Carnap was a logical positivist, holding roughly that all meaningful claims can be translated into claims about observable, verifiable phenomena. In the case of the mind, this comes to the claim that meaningful claims about the mind can be translated into claims about behavior. This is a form of *logical behaviorism*, holding ultimately that that what we *mean* when we make claims about the mind involves underlying claims about behavior. (This differs from scientific behaviorism, which holds roughly that the scientific study of the mind is the study of behavior.) Given that behavior itself seems to be a physical phenomenon, behaviorism can be seen as a form of materialism.

Like dualism, behaviorism has been subject to a number of objections. It seems more intuitive to say that mind is an inner cause of behavior, rather than an aspect of behavior itself. More concretely, one can argue that any given mental state is distinct from any given behavioral state or behavioral disposition. This sort of argument is mounted by Hilary Putnam in "Brains and Behavior" (chapter 7). Putnam argues that specially trained *beings* ("super-spartans") might feel pain while having no associated behavioral dispositions at all. Likewise, it can be argued that a perfect actor might have any given behavioral disposition without the associated mental state. If so, mental states cannot be behavioral dispositions.

### FURTHER READING

Apart from being the classic statement of a broadly behaviorist view, Ryle (1949) contains nuanced discussions of many aspects of the mind and mentality. The views of Hempel (1949), Wittgenstein (1953), and Quine (1960) also have some affinity with behaviorism. Dennett (chapter 52), who was a student of Ryle's, puts forward a view that

can be seen as a sophisticated contemporary descendant of behaviorism. Important objections to behaviorism are given by Geach (1957) and Block (1981). Scientific behaviorism is advocated by Watson (1930) and Skinner (1971).

Block, N. 1981. Psychologism and behaviorism. *Philosophical Review* 90:5–43.

Dennett, D. C. 1987. *The Intentional Stance*. MIT Press.

Geach, P. 1957. *Mental Acts*. Routledge and Kegan Paul.

Hempel, C. 1949. The logical analysis of psychology. In H. Feigl and W. Sellars, eds., *Readings in Philosophical Analysis*, pp. 373–84. Appleton-Century-Crofts.

Quine, W. V. 1960. *Word and Object*. MIT Press.

Ryle, G. 1949. *The Concept of Mind*. Hutchinson and Co.

Skinner, B. F. 1971. *Beyond Freedom and Dignity*. Alfred A. Knopf.

Watson, J. 1930. *Behaviorism*. Norton.

Wittgenstein, L. 1953. *Philosophical Investigations*. Blackwell.

### C. The Identity Theory

The mind–brain identity theory holds that mental states are brain states. Most theorists accept that mental states are at least associated or correlated with brain states: For example, feeling pain might be correlated with a certain sort of brain activity. The identity theory goes further to hold that mental states are *identical* to the associated brain states: These states are one and the same. This identification, unlike the behaviorist thesis, is not intended to be grounded in an analysis of our concepts. Rather, it is supposed to be an empirical claim, analogous to the claim that lightning is identical to electrical discharge or that water is identical to H<sub>2</sub>O. In this way, the identity theory can be seen as driven by scientific developments, especially in neuroscience.

The identity theory in its modern form was put forward by U. T. Place, J. J. C. Smart, and Herbert Feigl in the late 1950s. The original statement of the view was given by Place (chapter 8) and was refined and elaborated by his colleague Smart (chapter 9). Both Place and Smart recognize the strong intuitive resistance to the claim that mental states are brain states, especially in the case of conscious experiences. They try to defuse this resistance by careful diagnoses of its source: Place locates this in a “phenomenological fallacy,” while Smart addresses a number of objections to the thesis, putting forward the idea that mental concepts can be analyzed in a “topic-neutral way,” so that nothing in these concepts alone dictates whether or not mental states are physical. This suggests that while we it may not seem antecedently that mental states are physical, we can discover their physical nature through empirical science.

Feigl’s form of the identity theory has a somewhat different flavor from Place’s and Smart’s. Feigl gives more weight to intuitions about the special nature of conscious experience than do Place and Smart, but he argues that these can be reconciled with an identity view once we correctly understand the nature of physical understanding. Physical theory characterizes its entities structurally and leaves open their intrinsic nature. This raises the possibility that mental states are tied to the intrinsic aspect of physical states. This possibility is explored in the selection from Feigl here (chapter 10), which is a short excerpt from his article “The ‘Mental’ and the ‘Physical.’”

An objection to the identity theory, developed by Putnam (chapter 11), is that states such as pain cannot be identical to any particular brain states, since a creature such as a Martian might have pain without having the brain state in question. As Putnam put it, it is plausible that mental states are *multiply realizable*. If this is right, one cannot identify a mental state *type*, such as being in pain, with a physical state type, such as a specific sort of brain state. This still leaves open the possibility that one can identify mental state *tokens*, such as a specific pain of a subject, with physical state tokens, such as a specific

biological state of that subject. Many philosophers think that the type identity theory is refuted by this objection but that the token identity theory is left open.

### FURTHER READING

Feigl's long and interesting 1958 article is reprinted with an afterword as Feigl (1967). The identity theory provoked a great deal of critical discussion in the 1960s, some of which is collected in Borst (1970); another collection is Presley (1967). Influential objections to the identity theory are developed by Putnam (chapter 11) and Kripke (chapter 32). Hill (chapter 33; see also Hill 1991) advocates a version of the type identity theory.

Borst, C. V. (ed.). 1970. *The Mind/Brain Identity Theory*. Macmillan.

Feigl, H. 1967. *The 'Mental' and the 'Physical.'* University of Minnesota Press.

Hill, C. S. 1991. *Sensations: A Defense of Type Materialism*. Cambridge University Press.

Presley, C. P. (ed.). 1967. *The Identity Theory of Mind*. University of Queensland Press.

## D. Functionalism

Functionalism is a descendant of both behaviorism and the identity theory. Broadly speaking, it holds that mental states correspond to *functional* states: states of playing a certain role within the cognitive system. Some forms of functionalism identify mental states with functional states, while other forms identify mental states with the physical states that play the functional role in question. Functionalism was developed in the 1960s by Hilary Putnam, David Armstrong, and David Lewis, in somewhat different forms.

In Putnam's "The Nature of Mental States" (chapter 11), he argues against behaviorism and the identity theory and proposes instead a hypothesis that has come to be known as *machine functionalism*, according to which mental states are functional states of a computational machine. This makes mental states more abstract than any particular biological state, and so allows the possibility of multiple realization. It also allows a loose tie between mental states and behavior, without an absolute tie. This analogy between minds and machines has been very influential in contemporary philosophy of mind and cognitive science and has been developed in many different directions.

Armstrong (chapter 12) puts forward a different form of functionalism, cast not in terms of machines, but in terms of the general idea that mental states are defined in terms of their causal role. Specifically, he holds that the concept of a mental state is the concept of a state that is apt to be the cause of certain effects or apt to be the effect of certain causes. Where Putnam viewed his thesis as a sort of empirical hypothesis (like the identity theory), Armstrong puts forward his thesis as a sort of conceptual analysis (like logical behaviorism). It is a view about what we mean when we talk about the mind. Because of this, the view is often known as *analytic functionalism*. Unlike Putnam, Armstrong sees his view as supporting the identity theory rather than competing with it. If it turns out that in humans, a specific brain state plays the causal role associated with pain, then that brain state is itself a pain.

This sort of functionalism is developed in much more detail by Lewis (chapter 13). One might ask: *Which* causal roles are relevant to defining mental states? Lewis suggests that these roles are given by our everyday *theory* of the mind, as reflected in common-sense claims about mental states, their connections to one another, and their role in guiding behavior. If these claims are gathered together, they can be seen as providing a sort of definition of what it is to be the mental state in question: So pain is whatever state plays the role that common sense associates with pain, and so on. For this purpose, Lewis adapts Frank Ramsey's idea that the claims of a theory can be represented in a sin-

gle long sentence (the “Ramsey sentence”), and then used to identify the entities that satisfy the theory. Like Armstrong’s view, Lewis’ functionalism is a form of analytic functionalism, put forward as an analysis of what mental concepts mean.

A common objection to functionalism holds that it cannot deal with the “qualitative” aspects of conscious experience, such as the experience of seeing red or feeling pain. Ned Block (chapter 14) develops the “absent qualia” objection, according to which a system could have the same functional states as a conscious system, while having no qualitative states at all. He argues for this thesis using a thought-experiment involving a vast number of people causally organized to realize the given organization. Martine Nida-Rümelin (chapter 15) develops a version of the “inverted qualia” objection, according to which two systems could have the same relevant functional states as a conscious system while having *different* qualitative states. If these objections are correct, then qualitative states are not identical to functional states, so functionalism is false.

### FURTHER READING

There is an enormous and scattered literature on functionalism; some relevant papers are collected in Biro and Shahan (1982). Putnam’s version of functionalism is developed in a series of papers collected in Putnam (1975) and is repudiated in Putnam (1987). Armstrong’s functionalism is developed at length in Armstrong (1968). The idea that mental states are defined in terms of a theory is also present in Sellars (chapter 50) and is developed further by Churchland (chapter 53). Searle (chapter 63) gives an argument against machine functionalism that is closely related to the absent qualia argument. Dennett (chapter 26) tries to deflate the idea of absent and inverted qualia, while Shoemaker (1975) and White (1986) give important defenses of functionalism against absent qualia objections. See also a number of related papers on this topic (e.g., those by Nagel, Chalmers, Kripke, and Hill) in Part 2 of this book. Shoemaker (1982) and Palmer (1999) discuss inverted qualia from a philosophical and empirical standpoint, respectively.

- Armstrong, D. M. 1968. *A Materialist Theory of the Mind*. Routledge and Kegan Paul.
- Biro, J. I., and Shahan, R. W. (eds.). 1982. *Mind, Brain and Function*. Oklahoma University Press.
- Palmer, S. 1999. Color, consciousness, and the isomorphism constraint. *Behavioral and Brain Sciences* 22:1–21.
- Putnam, H. 1975. *Mind, Language, and Reality*. Cambridge University Press.
- Putnam, H. 1987. *Representation and Reality*. MIT Press.
- Shoemaker, S. 1975. Functionalism and qualia. *Philosophical Studies* 27:291–315. Reprinted in *Identity, Cause, and Mind* (Cambridge University Press, 1984).
- Shoemaker, S. 1982. The inverted spectrum. *Journal of Philosophy* 79:357–81. Reprinted in *Identity, Cause, and Mind* (Cambridge University Press, 1984).
- White, S. 1986. Curse of the qualia. *Synthese* 68:333–68. Reprinted in N. Block, O. Flanagan, and G. Güzeldere (eds.), *The Nature of Consciousness* (MIT Press, 1997).

## E. Other Psychophysical Relations

A number of other views about the relationship between the mental and the physical have been put forward. One important view is the *emergentism* of C. D. Broad (chapter 16) and others, holding that the mental is an emergent property of an underlying physical substrate. A property is *emergent* in Broad’s sense when it is a property of an underlying physical substance but cannot be deduced in principle from the low-level physical properties of that substance. Broad illustrates emergence in the domains of chemistry and biology, contrasting emergentism with *mechanism*, on which high-level properties are deducible from low-level properties, and with *substantial vitalism*, on which sepa-



rate substances are involved. Broad argues both for emergent qualities (high-level qualities that cannot be deduced from underlying qualities) and for emergent behavior (high-level behavior that cannot be deduced from the principles governing low-level behavior). This view of chemistry and biology is now widely rejected, although many theorists embrace a sort of *weak emergence* where high-level properties are surprising consequences of low-level properties, while still (unlike Broad's *strong emergence*) being deducible from them in principle. In the case of the mind (to which Broad applies the framework in a later chapter), strong emergence is more popular. In this domain, Broad's view has something in common with the interactionism of Descartes, but with just one underlying substance and two sorts of properties.

Another important view is the *anomalous monism* of Donald Davidson (chapter 17). This view can be seen as an attempt to preserve materialism without any strong reduction of the mental to the physical. On Davidson's view, any given mental event is identical to a physical event (a form of token identity theory), but there are no strict laws that connect mental events to physical events, and there are no strict laws governing mental events themselves. Davidson argues for this view by considering the distinct character of mental and physical concepts along with the causal connections between mental events and physical events.

There has been much discussion of whether the mental can be *reduced* to the physical, where this is understood as requiring more than the mere truth of materialism. Jerry Fodor (chapter 18) argues that in general, one cannot expect that the theories of a high-level "special science" should be reducible to the theories of a low-level science such as physics. Because of the many ways in which a high-level kind can be realized at a low level, the general principles in a high-level science cannot be captured by a low-level science except in a very complex and arbitrary way. This applies especially to the science of psychology, suggesting that one cannot expect that psychology can be reducible to physics, or even to neuroscience. Instead, it will always have a degree of autonomy.

In a more recent paper (chapter 19), Jaegwon Kim replies to Fodor's argument, suggesting that multiple realizability does not pose an obstacle to reduction. We may still be able to reduce the psychological principles that apply to *humans* (or some other particular species) to lower-level principles and to reduce species-specific psychological kinds to lower-level kinds. One should not expect anything more, since there may not be any general principles that apply to all possible minds. Where one has multiple realizability, one has multiple psychological principles. So the important principles and kinds always lie at a species-specific level.

It has also been suggested that the mental-physical relationship can be analyzed using the notion of *supervenience*. A class of high-level properties supervenes on a class of low-level properties when any two possible systems (or possible worlds) that have the same low-level properties have the same high-level properties. This captures the intuitive materialist idea that a system's physical properties might determine its mental properties, without making further commitments about the relationship. The idea is discussed critically by Terence Horgan (chapter 20). Horgan sketches the history and the various versions of supervenience and argues that while supervenience is a useful tool, it needs to be augmented to yield a full materialist theory of the mental-physical relations. In particular, we need an *explanation* of why the mental properties supervene on physical properties. Supervenience augmented by this sort of explanation yields what Horgan calls *superdupervenience*.

Finally, Frank Jackson (chapter 21) argues for an important role for *conceptual analysis* in understanding the mental-physical relation. Jackson argues that the truth of materialism requires a sort of supervenience of the mental on the physical and he argues that this in turn requires that mental concepts can be analyzed in such a way so that there is a sort of a priori entailment from physical truths to mental truths. This sort of analyzability can be seen as one way of answering Horgan's request for an explanation of superve-