



*Mind in a
Physical World*

An Essay on the Mind-Body Problem
and Mental Causation

Jaegwon Kim

Mind in a Physical World

Representation and Mind

Hilary Putnam and Ned Block, editors

Representation and Reality by Hilary Putnam

Explaining Behavior: Reasons in a World of Causes by Fred Dretske

The Metaphysics of Meaning by Jerrold J. Katz

A Theory of Content and Other Essays by Jerry A. Fodor

The Realistic Spirit: Wittgenstein, Philosophy, and the Mind by Cora Diamond

The Unity of the Self by Stephen L. White

The Imagery Debate by Michael Tye

A Study of Concepts by Christopher Peacocke

The Rediscovery of the Mind by John R. Searle

Past, Space, and Self by John Campbell

Mental Reality by Galen Strawson

Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind
by Michael Tye

Representations, Targets, and Attitudes by Robert Cummins

Starmaking: Realism, Anti-Realism, and Irrealism edited by Peter J. McCormick

A Logical Journey: From Gödel to Philosophy by Hao Wang

Brainchildren: Essays on Designing Minds by Daniel C. Dennett

Realistic Rationalism by Jerrold J. Katz

The Paradox of Self-Consciousness by José Luis Bermúdez

In Critical Condition: Polemical Essays on Cognitive Science and the Philosophy of Mind by Jerry Fodor

Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation by Jaegwon Kim

Mind in a Physical World

An Essay on the Mind-Body Problem and Mental
Causation

Jaegwon Kim

A Bradford Book
The MIT Press
Cambridge, Massachusetts

First MIT Press paperback edition, 2000
©1998 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form or by any electronic or mechanical means (including information storage and retrieval) without permission in writing from the publisher.

This book was set in Palatino by Wellington Graphics

Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Kim, Jaegwon.

Mind in a physical world : an essay on the mind-body problem and mental causation / Jaegwon Kim.

p. cm. — (Representation and mind)

“A Bradford book.”

Book consists of four revised lectures given by the author as the Townsend lectures at the University of California, Berkeley, in Mar. 1996.

Includes bibliographical references and index.

ISBN 0-262-11234-5 (hardcover : alk. paper), 0-262-61153-8 (pb)

1. Philosophy of mind. 2. Mind and body. 3. Causation.

I. Title. II. Series.

BD418.3.K53 1998

128'.2—dc21

98-24346

CIP

Contents

Preface vii

Chapter 1

The Mind-Body Problem: Where We Now Are 1

Supervenience, Realization, and Emergence 4

Supervenience Is Not a Mind-Body Theory 9

The Layered Model and Mereological Supervenience 15

Physical Realizationism 19

Physical Realizationism Explains Mind-Body Supervenience 23

Chapter 2

The Many Problems of Mental Causation 29

Three Problems of Mental Causation 32

The Supervenience Argument, or Descartes's Revenge 38

Searle, Fodor, and the Supervenience Argument 47

Block's Worries about Second-Order Properties 51

Chapter 3

Mental Causation: The Backlash and Free Lunches 57

Unavoidability of Metaphysics: The Exclusion Problem 60

Do Counterfactuals Help? 67

"Program Explanation" and Supervenient Causation 72

Does the Problem of Mental Causation Generalize? 77

Properties: "Levels" and "Orders" 80

Chapter 4

Reduction and Reductionism: A New Look 89

Nagel Reduction: Troubles with "Bridge Laws" 90

The Functional Model of Reduction 97

vi Contents

<i>Functional Properties versus Functional Concepts</i>	103
<i>Multiple Realization Again</i>	106
<i>The Supervenience Argument Revisited</i>	112
<i>The Options: Good News and Bad News</i>	118
Notes	121
References	135
Index	141

Preface

This book consists of the four lectures I gave as the Townsend Lectures at University of California, Berkeley, in March, 1996. The texts have been extensively revised since then, but there have been no significant changes in the main points of my claims and arguments. My thanks go to the Berkeley Philosophy Department for the invitation, and for their hospitality during my visit. I benefited from comments and questions from members of the audience, in particular, David Chalmers, Martin Jones, and David Sosa.

In these lectures I tried to set out my current thoughts on a range of issues concerning the metaphysics of the mind—in particular, the mind-body problem, mental causation, and reductionism. The texts were originally written with oral presentation in mind, and I have decided to retain in this final version their somewhat informal and relaxed style of presentation.

These lectures draw materials from the following recent papers of mine: “Mental Causation: What? Me Worry?,” *Philosophical Issues* 6 (1995): 123–151; “The Mind-Body Problem: Taking Stock After 40 Years,” *Philosophical Perspectives*, 1997; “What is the Problem of Mental Causation?,” *Norms and Structures in Science*, ed. M. L. Dalla Chiara et al. (Dordrecht: Kluwer, 1997); “Does the Problem of Mental Causation Generalize?,” *Proceedings of the Aristotelian Society*, 1997.

In thinking about the issues discussed here I am indebted to many friends and colleagues who share my interest in philosophy of mind. Who many of them are should be obvious from my texts and notes. I also benefited from spirited comments and challenges from my students at Brown. Maura Geisser, my

assistant, has been extremely helpful with her dedicated and efficient work in preparing the manuscript, proofreading, and other tedious but necessary chores. I also want to thank Brown University for its support of my work over the years and my departmental colleagues and students for providing me with a friendly and stimulating environment in which to work.

Providence, Rhode Island
April 1997

Mind in a Physical World

Chapter 1

The Mind-Body Problem: Where We Now Are

Current debates on the mind-body problem can be traced back to the late 1950s and early 1960s. To be more precise, arguably the mind-body problem as we now know it had its origin in two classic papers published one year apart: "The 'Mental' and the 'Physical'" by Herbert Feigl in 1958 and "Sensations and Brain Processes" by J. J. C. Smart the following year.¹ In these papers Smart and Feigl independently proposed an approach to the nature of mind that has come to be called the mind-body identity theory, central-state materialism, the brain state theory, or type physicalism. Although U. T. Place's "Is Consciousness a Brain Process?,"² published in 1956, anticipated Smart and Feigl, it was the papers by Smart and Feigl that reintroduced the mind-body problem as a mainstream metaphysical Problematik of analytical philosophy, and launched the debate that has continued to this day. True, Ryle's *The Concept of Mind* was out in 1948, and there were of course Wittgenstein's much debated remarks on mentality and mental language, not to mention a much earlier work by C. D. Broad, *The Mind and Its Place In Nature* (1925).³ But Ryle's and Wittgenstein's primary concerns were directed at the "logic" of mental discourse rather than the metaphysical problem of explaining how our mentality is related to our physical nature, and moreover Ryle and Wittgenstein, each for different reasons, would have denounced the metaphysical mind-body problem as a piece of philosophical nonsense. In contrast, Broad's work was robustly metaphysical, but unfortunately, it failed to connect with the mind-body debate in the second half of this century, especially in its important early stages.

For many of us who, like me, went to graduate school in the late 1950s and early 1960s, Smart's and Feigl's materialism was our first encounter with the mind-body problem as a systematic philosophical problem. Their approach sounded refreshingly bold and tough-minded, and seemed in tune with the optimistic scientific temper of the times. It was an intriguing and exciting idea that mental events could just *be* brain processes, and that scientific research could show this, just as science showed us that light was electromagnetic radiation, and that genes were DNA molecules. But the identity theory was unexpectedly short-lived—its precipitous fall began only several years after its introduction. It is clear in retrospect, though, that despite its short life, the theory made one crucial contribution that has outlasted its reign as a theory of the mind. What I have in mind is the fact that the brain state theory helped set the basic parameters and constraints for the debates that were to come—a set of broadly physicalist assumptions and aspirations that still guide and constrain our thinking today. One indication of this is the fact that when the brain state theory began fading away in the late 1960s and early 1970s few lapsed back into Cartesianism or other serious forms of mind-body dualism. Almost all the participants in the debate have stayed with physicalism, and even those who had a major hand in the demise of the Smart-Feigl materialism have continued their allegiance to a physicalist worldview. Through the 1970s and 1980s and down to this day, the mind-body problem—*our* mind-body problem—has been that of finding a place for the mind in a world that is fundamentally physical. The shared project of the majority of those who have worked on the mind-body problem over the past few decades has been to find a way of accommodating the mental within a principled physicalist scheme, while at the same time preserving it as something distinctive—that is, without losing what we value, or find special, in our nature as creatures with minds.

What made the demise of the mind-brain identity theory so quick and seemingly painless, causing few regrets and second thoughts among philosophers, was the fact the two principal objections that overthrew it, the multiple realization argument

advanced by Hilary Putnam⁴ and the anomalist argument of Donald Davidson,⁵ contained within them seeds for appealing alternative pictures of mentality, namely functionalism and anomalous monism. The core idea of functionalism, that mental kinds and properties are functional kinds at a higher level of abstraction than physicochemical or biological kinds, was a suggestive and eye-opening idea that seemed to help us make sense of cognitive science, which was being launched about the same time. The functionalist approach to mentality seemed made to order for the new science of mentality and cognition, for its central doctrine seemed to postulate a distinctive domain of mental/cognitive properties that could be scientifically investigated independently of their physical/biological implementations—an idea that promised for psychology both legitimacy and autonomy as a science. Functionalism provided the new science of cognition with both a metaphysics and a methodology.

Davidson's anomalous monism, too, offered us another appealing package, although its special attractions were different from those of functionalism. On the one hand, it told us that the mental domain, on account of its essential anomalousness and normativity, cannot be the object of serious scientific investigation,⁶ placing the mental on a wholly different plane from the physical. In particular, its anomalousness *vis-à-vis* the physical—that is, the impossibility of laws connecting mental kinds with physical kinds—was thought to entail the irreducibility of mental kinds to physical kinds. This meant that mental kinds are distinct from physical or biological kinds, contradicting the Smart-Feigl mind-body identity thesis. In this way Davidson's anomalous monism secured for us the autonomy of the mental, although for reasons that are different from those offered by functionalism. On the other hand, the monistic component of anomalous monism insisted that all individual events ("token events") are physical events subject to laws of physics, thereby giving the physical the pride of place in our ontology. This assuaged our physicalist aspirations. Each in its own distinctive way, therefore, both anomalous monism and functionalism made it possible for us to shed the restrictive constraints of

monolithic reductionism without losing our credentials as physicalists. Or so it seemed.

Supervenience, Realization, and Emergence

What do these two doctrines, anomalous monism and functionalism, say about the mind-body problem, the problem of how the mental and the physical are related? Anomalous monism is a physicalist monism that holds that every individual mental event is a physical event, but it also maintains that the mental is anomalous, that is, not governed by laws (or “strict laws,” as Davidson sometimes puts it). In particular, it famously claims that there are no laws connecting mental kinds or properties with physical ones. This component of the doctrine, the anomalousness of the mental in relation to the physical, is a *negative* thesis: it tells us how the mental is *not* related to the physical, and says nothing about how the two *are* related. The burden therefore falls on the first component of the doctrine—the claim that every mental event is a physical event—to tell us a positive story about the relationship between the mental and the physical.

But exactly what does Davidson’s physical monism tell us about the mind-body relation? The answer is: much less than what one might have expected. For Davidson, the content of the claim that all mental events are physical events turns out to be only this: every event that can be given a mental description can also be given a physical description, or, as we might say in current idiom, that every event that has a mental property (or falls under a mental kind) also has a physical property (falls under a physical kind). For, within Davidson’s scheme, an event is physical or mental only as it is describable in the physical or mental vocabulary, or as it falls under a physical or a mental kind. So Davidson’s monism comes to this proposition: there are no events that have only mental properties (descriptions), although there may be, and presumably are, events with physical properties (descriptions) only. This requires no type-type connections between mental and physical kinds, and Davidson’s

doctrine of mental anomalism specifically prohibits nomological (or any stronger) linkages between mental types and physical types. But a doctrine that outlaws type-type, or property-property, connections between the mental and the physical is one that in effect says there is no connection between our mental nature and our physical nature. This means that the monism of anomalous monism is no less a negative thesis than the anomalism of anomalous monism.

That anomalous monism doesn't say much about the mind-body relation can be seen from an analogy. Consider the statement, which presumably is true, that every object that has a color has a shape—or to put it in a way that is parallel to Davidson's monism, every object with a color is identical with an object with a shape. Clearly this statement says nothing about the relation between colors and shapes; in fact we know there are no interesting connections between them. Analogously in the mind-body case, the statement that every event with a mental property is an event with a physical property says nothing about—and it is designed by Davidson to say nothing about—the presence or absence of type-type connections between the mental and the physical. This means that Davidson's anomalous monism says no more about the relationship between the mental and the physical than the claim that all objects with a color have a shape says about the relationship between colors and shapes. As far as anomalous monism goes then, there need be no systematic relationships between mental properties and physical properties—any more than between the colors and shapes of things around us. The same holds for any version of the so-called token physicalism similar to Davidson's physical monism.⁷ I believe we want our mind-body theories to tell us more, a *positive* story about how mental properties and physical properties *are* related, and hopefully also explain why they are so related. We don't get such a story from anomalous monism.

This may well have been what led Davidson, in his "Mental Events," to invoke mind-body supervenience. After developing his argument for the impossibility of laws connecting the mental and the physical, he adds in a somewhat offhanded way:

Although the position I describe denies there are psychophysical laws, it is consistent with the view that mental characteristics are in some sense dependent, or supervenient, on physical characteristics. Such supervenience might be taken to mean that there cannot be two events alike in all physical respects but differing in some mental respects, or that an object cannot alter in some mental respect without altering in some physical respect.⁸

Although the British emergentists early in this century appear to have been the first to use the expression “supervenience” in connection with the mind-body problem and the concept had been around for some time in ethical theory, the passage just cited signaled the introduction of the supervenience idiom into the contemporary debate on the mind-body problem. In any case mind-body supervenience understood in some such sense as Davidson’s quickly caught on, I believe, precisely because it gave, or at least promised to give, a statement of how the mental and the physical *are* related to each other. Note, in particular, the fact that supervenience concerns mental and physical *properties* and *kinds* (or “characteristics” or “respects,” in Davidson’s words), not mental and physical tokens or Davidsonian events as structureless particulars. Unlike the two central theses of anomalous monism, the supervenience claim therefore finally says something positive about the relationship between mental and physical properties.

Moreover the supervenience claim is physicalistically appealing: an asymmetric dependence of the mental on the physical is clearly implied, if not stated outright. This mind-body dependence relation was thought (in fact, explicitly claimed by Davidson and others) to be consistent with the irreducibility of the mental to the physical. It is fair to say, I think, that the supervenience idea, which Davidson seemed to toss off almost like an afterthought and which he did little to explain or elaborate, much less defend, came to overshadow the core doctrines of anomalous monism, creating a new focus and direction for the physicalist program. Whether mind-body supervenience is to be

thought of as an integral component of Davidson's anomalous monism is not a question that concerns us here.⁹ Whatever we may want to say about that question, the fact remains that philosophers found a promising physicalist metaphysics in mind-body supervenience rather than in the two theses of anomalous monism, that is, physicalist event monism and the anomalism of the mental. That is why the supervenience idea took off, in the late 1970s, with a life of its own in the mind-body debate.

Functionalists, by and large, were not metaphysicians, and few of them were self-consciously concerned about just what their position entailed about the mind-body problem.¹⁰ The key term the mainstream functionalists used to describe the relation between mental properties (kinds, states, etc.) and physical properties was "realization" (or sometimes "implementation," "execution," etc.): mental properties are "realized" or "implemented" by (or in) physical properties, though neither identical nor reducible to them. But the term "realization" was introduced,¹¹ and quickly gained currency, chiefly on the strength of computational analogies (in particular, the idea that abstract, mathematically characterized computing machines are realized in concrete physical/electronic devices), and few functionalists, especially in the early days, made an explicit effort to explain what the realization relation consisted in—in particular, what this relation implied as regards the traditional options on the mind-body problem.

I believe that the idea of supervenience was embraced by some functionalists, and those influenced by the multiple realization argument, in part because it promised to fill this metaphysical void. The thesis that mental properties are supervenient on physical properties seemed to fit the metaphysical requirements of functionalism very nicely: it promised to give a clear sense to the primacy of the physical domain and its laws, thereby doing right by the physicalist commitments of most functionalists—and this without implying physical reductionism, thereby protecting the autonomy of the mental. Further, in allowing multiple supervenience bases for supervenient properties, it

offered a perfect scheme for accommodating the multiple realizability of mental properties.

I believe this is why many philosophers, including those who espoused the functionalist outlook, saw in mind-body supervenience a satisfying metaphysical statement of physicalism without reductionism. This helped to create, in the mid- to late 1970s, what Ned Block has called “the antireductionist consensus”¹² and its subsequent entrenchment. This position, now standardly called “nonreductive physicalism” (or “nonreductive materialism”), has been, and still is, the most influential metaphysical position, not only on the mind-body problem but more generally on the relationship between higher-level properties and underlying lower-level properties in all areas. Thus the approach yielded as a bonus a general philosophical account of how the special sciences are related to basic physics: properties in the domain of the special sciences—properties in terms of which laws and explanations in these sciences are formulated—are supervenient on basic physical properties but not reducible to them, and in this sense the special sciences are autonomous from basic physics. If you are a scientist outside basic physics, it is probably comforting to know that you can do your science independently of what goes on in theoretical physics, and that you can formulate laws and explanations in your special discipline in its own distinctive vocabulary that need not answer to that of physics. Of course this is what scientists outside basic physics generally do anyway, but it’s nice to have a philosophical rationale for it.

In any case, one direct consequence of the entrenchment of the antireductionist consensus has been the return of emergentism—if not all the florid doctrines of classic emergentism of the 1920s and 1930s, at least its distinctive vocabulary and slogans. In the heyday of positivism and “unity of science,” emergentism used to be relegated to the heap of unsavory pseudoscientific doctrines, not quite as disreputable as, say, neo-vitalism with its entelechies and *élan vital*, but nearly as obscure and incoherent. With the demise of reductive physicalism, emergentism has been showing strong signs of a comeback.¹³ We now see an increasing

and unapologetic use of expressions like “emergent,” “emergent property,” and “emergent phenomenon,” seemingly in the sense intended by the classic emergentists, not only in serious philosophical literature¹⁴ but in the writings in psychology, cognitive science, systems theory, and the like.¹⁵

To sum up, then, three ideas have been, and still are, prominent in discussions of the mind-body problem since the demise of the reductive brain-state theory: the idea that the mental “supervenes” on the physical, the idea that the mental is “realized” by the physical, and the idea that the mental is “emergent” from the physical. Exploration of these three ideas, and the roles they play in debates on the mind-body problem, will be one of the things I want to do in these lectures. Today I will focus on supervenience and realization. I will say something about emergence along the way as my talks proceed, and especially in connection with my discussion of reduction and reductionism in a later talk (chapter 4).¹⁶

Supervenience Is Not a Mind-Body Theory

Let us begin with supervenience. Supervenience is standardly taken as a relation between two sets of properties, the supervenient properties and their base properties. As is by now well known, a variety of supervenience relations is available; for our purposes we may focus on what is called “strong supervenience.” Consider then the following thesis of mind-body supervenience:

Mental properties *supervene* on physical properties, in that necessarily, for any mental property *M*, if anything has *M* at time *t*, there exists a physical base (or subvenient) property *P* such that it has *P* at *t*, and necessarily anything that has *P* at a time has *M* at that time.

For example, if a person experiences pain, it must be the case that that person instantiates some physical property (presumably, a complex neural property) such that whenever anyone instantiates this physical property, she must experience pain.

That is, every mental property has a physical base that guarantees its instantiation. Moreover, without such a physical base, a mental property cannot be instantiated. Under certain assumptions concerning property composition, a supervenience thesis stated this way (sometimes called “the modal operator” formulation) can be shown to be equivalent to another familiar formulation of supervenience (sometimes called “the possible world” or “indiscernibility” definition):

Mental properties *supervene* on physical properties, in that necessarily any two things (in the same or different possible worlds) indiscernible in all physical properties are indiscernible in mental respects.

Or we can say: any two things that are exact physical duplicates necessarily are exact psychological duplicates as well—that is, physical duplicates are duplicates *tout court*. Or, as some have put it: no mental difference without a physical difference.¹⁷ We will consider these two formulations of mind-body supervenience to be equivalent¹⁸ and use one or the other to suit the context.

Under mind-body supervenience a physical base property, *P*, for mental property *M* guarantees, as a matter of necessity, the occurrence of *M*; that is, necessarily if something instantiates *P*, it instantiates *M*. The modal force of necessity involved is a parameter to be fixed to suit one’s view of the mind-body relation; some may go for metaphysical or even logical/conceptual necessity, while others will settle for nomological necessity. (We should keep in mind that the modal force of supervenience may vary for different groups of mental properties; for example, it is possible for intentional properties to supervene with logical/conceptual necessity, while phenomenal properties supervene only with nomological necessity.) As has been noted, one and the same mental property may have multiple physical bases: an instance of pain in a human may be grounded in one neural property, and another instance of pain, say in a reptile, may be grounded in another.

As we saw in Davidson, it is customary to associate supervenience with the idea of *dependence* or *determination*: if the mental supervenes on the physical, the mental is dependent on the physical, or the physical determines the mental, roughly in the sense that the mental nature of a thing is entirely fixed by its physical nature. Sometimes this is put in terms of “worlds”: the psychological character of a world is determined entirely by its physical character—as it is often put, worlds that are physically indiscernible are psychologically indiscernible. The relation of dependence, or determination, is asymmetric: if x depends on, or is determined by y , it cannot be that y in turn depends on or is determined by x . What does the determining must be taken to be, in some sense, ontologically prior to, or more basic than, what gets determined by it. But mind-body supervenience as stated isn’t asymmetric; in general, the supervenience of A on B does not exclude the supervenience of B on A . The notion of supervenience we introduced simply states a pattern of *covariance* between the two families of properties, and such covariances can occur in the absence of a metaphysical dependence or determination relation. For example, two sets of properties may show the required covariance because each depends on a third, somewhat in the manner in which two collateral effects of the single cause exhibit a pattern of lawful correlation. What needs to be added to property covariance to get dependence or determination, or whether dependence/determination must be taken as an independent primitive, are difficult questions that probably have no clear answers. We will simply follow the customary usage and understand supervenience to incorporate a dependence/determination component as well. In fact common expressions like “supervenience base” and “base property” all but explicitly suggest asymmetric dependence.

Suppose then that the mental supervenes on the physical. Does this give us a possible account of how our mentality is related to the physical nature of our being? That is, can we use supervenience itself to state a philosophical theory of the way minds are related to bodies? It has sometimes been thought—as

I myself did at one point—that the answer is yes, that what might be called supervenience physicalism is a possible position to take on the mind-body problem. There has been a controversy concerning whether supervenience, in the sense of strong supervenience, is indeed consistent with the irreducibility of the supervenient properties to their subvenient bases. But the discussion of this question has been inconclusive,¹⁹ and I now believe that the debate was framed in terms of a seriously flawed notion of reduction (see chapter 4). Here we will focus on the question whether or not mind-body supervenience as such can be thought of as an account of the mind-body relation.

Brief reflection shows that the answer is no, that mind-body supervenience in itself does not give us a theory of the mind-body relation. There are at least two related reasons for this. First, mind-body supervenience is consistent with a host of classic positions on the mind-body problem; in fact it is a shared commitment of many mutually exclusionary mind-body theories. As we will see, both emergentism and the view that the mental must be physically realized—that is, there can be no nonphysical realizations of mental properties (we can call this physical realizationism)—imply mind-body supervenience. But emergentism is a form of dualism that takes mental properties to be nonphysical intrinsic causal powers, whereas physical realizationism, as I will argue, is a monistic physicalism. What is more obvious, type physicalism, which reductively identifies mental properties with physical properties, implies mind-body supervenience. Moreover epiphenomenalism, viewed by some physicalists, for example, J. J. C. Smart, as their chief dualistic rival, is apparently committed to mind-body supervenience: if two organisms differ in some mental respect, that must be because they differ in some physical respect—it must be because the physical cause of the mental respect involved is present in one but absent from the other. That is, the epiphenomenalist would surely agree that two physically indistinguishable organisms must manifest the same mental characteristics. If mind-body supervenience is a commitment of each of these diverse, and conflicting, approaches to the mind-body problem, it cannot

itself be a position on this problem that can be set alongside these classic alternatives.²⁰

What this shows is that the mere claim of mind-body supervenience leaves unaddressed the question what *grounds* or *accounts for* it—that is, the question why the supervenience relation should hold for the mental and the physical.²¹ To see the general issue involved here, consider normative supervenience, the widely accepted doctrine that normative or evaluative properties supervene on nonnormative, nonevaluative properties. Various metaethical positions accept normative supervenience but offer differing accounts of its provenance. For the ethical naturalist, the supervenience holds because normative properties are definable in terms of nonnormative, naturalistic properties. An ethical intuitionist like G. E. Moore would see normative supervenience as a fundamental synthetic a priori fact not susceptible to further explanation; it is something we directly apprehend through our moral sense. R. M. Hare, a noncognitivist, would attempt to tie it with certain regulative constraints on the language of prescription. Still others may try to ground it in the very idea of normative evaluation: normative judgments and evaluations must ultimately be based on reasons or grounds that are themselves nonnormative and nonevaluative, and this means that normative properties must have nonnormative criteria of application. In the mind-body case, too, we can think of rival mind-body theories as offering competing explanations of mind-body supervenience. The explanation offered by reductive type physicalism is analogous to the naturalistic explanation of normative supervenience—mind-body supervenience holds because mentality is physically reducible. On type physicalism, mental properties *are* physical properties, just as on ethical naturalism ethical properties *are* natural properties. Emergentism, like ethical intuitionism, views mind-body supervenience as something that admits no explanation; it is a brute fact that must be accepted with “natural piety,” as a leading emergentist, Samuel Alexander, urged. In contrast, epiphenomenalism would invoke the causal relation (the “same cause, same effect” principle) to explain supervenience, while on physical realizationism

mind-body supervenience is a direct consequence of the fact that mental properties are, as we will see, second-order functional properties defined over first-order physical properties, and so on.

We must conclude then that mind-body supervenience itself is not an *explanatory theory*; it merely states a pattern of property covariation between the mental and the physical and points to the existence of a dependency relation between the two. Yet supervenience is silent on the nature of the dependence relation that might explain why the mental supervenes on the physical. Another way of putting the point would be this: supervenience is not a *type* of dependence relation—it is not a relation that can be placed alongside causal dependence, reductive dependence, mereological dependence, dependence grounded in definability or entailment, and the like. Rather, any of these dependence relations can generate the required covariation of properties and thereby qualify as a supervenience relation. Supervenience therefore is not a metaphysically “deep” relation; it is only a “phenomenological” relation about patterns of property covariation, patterns that possibly are manifestations of some deeper dependence relationships. If this is right, mind-body supervenience *states* the mind-body problem—it is not a solution to it. This means that nonreductive physicalism must look elsewhere for its metaphysical grounding; supervenience itself is not capable of supplying it. Any putative account of the mind-body relation that accepts mind-body supervenience must specify a dependence relation between the mental and physical that is capable of grounding and explaining mind-body supervenience.

These considerations, however, need not be taken to be seriously deflationary about the usefulness of the supervenience idea in the philosophy of mind. They certainly deflate the hope that supervenience itself might give us an account of the mind-body relation. But there is also a positive side: our considerations indicate that mind-body supervenience captures a commitment common to all positions on the nature of mentality that are basically physicalistic. For it represents the idea that mentality is at bottom physically based, and that there is no free-floating

mentality unanchored in the physical nature of objects and events in which it is manifested. This is an idea that can be shared by many diverse positions on the mind-body problem, from reductive type physicalism at one extreme to dualistic emergentism at the other. In contrast, mind-body supervenience is inconsistent with more extreme forms of dualism, such as, Cartesian dualism, which allow the mental world to float freely, unconstrained by the physical domain.²² Thus mind-body supervenience can serve as a useful dividing line: it can be viewed as defining *minimal physicalism*.

The Layered Model and Mereological Supervenience

Cartesian substance dualism pictures the world as consisting of two independent domains, the mental and the material, each with its own distinctive defining properties (consciousness and spatial extendedness, respectively). There are causal interactions across the domains, but entities in each domain, being “substances,” are ontologically independent of those of the other, and it is metaphysically possible for one domain to exist in the total absence of the other. What has replaced this picture of a dichotomized world is the familiar multilayered model that views the world as stratified into different “levels,” “orders,” or “tiers” organized in a hierarchical structure. The bottom level is usually thought to consist of elementary particles, or whatever our best physics is going to tell us are the basic bits of matter out of which all material things are composed.²³ As we go up the ladder, we successively encounter atoms, molecules, cells, larger living organisms, and so on. The ordering relation that generates the hierarchical structure is the mereological (part-whole) relation: entities belonging to a given level, except those at the very bottom, have an exhaustive decomposition, without remainder, into entities belonging to the lower levels. Entities at the bottom level have no physically significant proper parts.

It is part of this layered picture that at each level there are properties, activities, and functions that make their first appearance, or “emerge,” at that level (we may call them the

characteristic properties of that level). Thus among the characteristic properties of the molecular level are electrical conductivity, inflammability, density, viscosity, and the like; activities and functions like metabolism and reproduction are among the characteristic properties of the cellular and higher biological levels; and consciousness and other mental properties make their appearance at the level of higher organisms. For much of this century, a layered picture of the world like this has formed a constant—tacitly assumed if not explicitly stated—backdrop for debates on a variety of issues in metaphysics and philosophy of science—for example, reduction and reductionism, the mind-body problem, emergence, the status of the special sciences, and the possibility of a unified science. In fact this picture has had a strong and pervasive influence on the way we formulate problems and their possible solutions in many areas. Sometimes the layered model is couched in terms of concepts and languages rather than entities and their properties. Talk of levels of *organization*, *descriptions* or *languages*, of *analysis*, of *explanation*, and the like is encountered everywhere—it has thoroughly permeated primary scientific literature in many fields, in particular, various areas of psychology and cognitive science, systems theory, and computer science—as well as philosophical writings about science.²⁴

Now we come to a crucial question: How are the characteristic properties of a given level related to the properties at the adjacent levels—in particular, to those at the lower levels? How are biological (“vital”) properties related to physicochemical properties? How are consciousness and intentionality related to biological/physical properties? How are social phenomena, phenomena characteristic of social groups, related to phenomena involving individual members? As you will agree, these are among the central questions of philosophy of science, metaphysics, and philosophy of mind. Possible answers to them define the philosophical options on the issues involved. Some of the well-known major alternatives include reductionism, antireductionism, methodological individualism, emergentism, neo-vitalism, and the like. You may attempt to give a single uniform answer

applicable to all pairs of adjacent levels, or you may take different positions regarding different levels. For example, you might argue that properties at every level (higher than the bottom level) are reducible, in some clear and substantial sense, to lower-level properties and therefore ultimately to the basic properties of physics, or you might restrict the reductionist claim to certain selected levels (say, biological properties in relation to physicochemical properties) and defend an anti-reductionist stance concerning properties at other levels (say, consciousness and intentionality). And it isn't even necessary to give a uniform answer in regard to all characteristic properties of a given level; concerning mental properties, for example, it is possible to hold that phenomenal properties, or qualia, are irreducible, while holding intentional properties, including propositional attitudes, to be reducible (say, causal/functionally or biologically).

Let us now look at the layered model with supervenience in mind. When supervenience is superposed on the layered model, something like the following emerges as a general schema of supervenience claims about properties at a given level (other than the lowest one) in relation to those at lower levels:

For any x and y , belonging to level L (other than the lowest level), if x and y are indiscernible in relation to properties at all levels lower than L (or, as we may say, x and y are *microindiscernible*), then x and y are indiscernible with respect to all properties at level L .

How do we explain the idea of *microindiscernibility*? The following seems pretty natural and straightforward:²⁵

x and y , belonging to level L , are *microindiscernible* if and only if for every decomposition D of x into proper parts belonging to lower levels, y has an isomorphic decomposition C in the sense that there is a one-one function I from D to C such that for any n -adic property or relation P at levels lower than L , $P(\mathbf{d}_n)$ iff $P(I(\mathbf{d}_n))$, where \mathbf{d}_n is any n -tuple of elements in D and $I(\mathbf{d}_n)$ is the image of \mathbf{d}_n under I , and conversely from y to x .

Not surprisingly, supervenience theses, when applied to the layered model, turn into claims of *mereological supervenience*, the doctrine that properties of wholes are fixed by the properties and relations that characterize their parts. A general claim of macro-micro supervenience then becomes the Democritean atomistic doctrine that the world is the way it is because the microworld is the way it is.²⁶

Let us now return to mental properties. Presumably mental properties arise as characteristic properties at the level of higher organisms, and like any other high-level properties, they are supervenient in the sense explained, on the lower-level properties characterizing their proper parts. That is, if *M* is a mental property had by something *x*, then any *y* that is microindiscernible from *x* will also have *M*. Mental properties, therefore, are macroproperties supervening on microproperties.

We should resist the temptation to read more into this result than what's really there. On the layered model, mind-body supervenience is an instance of *mereological supervenience*, and this might seem like an advance, tempting us into thinking that we might try explaining mind-body supervenience in parallel with the way macrophysical properties are determined and explained by microphysical properties. But supervenience or determination is one thing, explanation quite another. We may know that *B* determines *A* (or *A* supervenes on *B*) without having any idea why this is so—why *A* should arise from *B*, not *C*, or why *A*, rather than *D*, arises from *B*. Mereological supervenience of the mental on the physical would not automatically promise us an intelligible account of why the particular mind-body supervenience relations hold. Given that mental property *M* is supervenient on a certain physical mereological configuration *P*, the questions still remain: Is *M* reducible to *P* in some appropriate sense? Can we explain why something has *M* in terms of its having *P*? Are the *P*-*M* and other such supervenience relations further explainable (and what can “explanation” mean here?), or must they be taken as brute and fundamental? These questions are independent of the question whether *P* is a microphysically characterized property.

These, I believe, are legitimate questions. The layered model provides us with a useful ontological scheme within which we can situate the mind-body problem. By placing the problem in the broader context of this model, we give it both generality and structure, and this raises the hope that we will be able to set some nonarbitrary parameters and constraints on the resolution of the problem, and to appreciate its possible distinctiveness as well as its commonality with problems in other areas. But the hard questions remain untouched. Let us now turn to the idea of physical realization as an approach to these questions.

Physical Realizationism

As you may recall, by physical realizationism I mean the claim that mental properties, if they are realized, must be physically realized—that is, no mental properties can have nonphysical realizations.²⁷ The thesis therefore is equivalent to the conjunction of physicalism with the functionalist conception of mental properties, and “physicalist functionalism” would be an equally good name for this position. Functionalism takes mental properties and kinds as functional properties, properties specified in terms of their roles as causal intermediaries between sensory inputs and behavioral outputs, and the physicalist form of functionalism takes physical properties as the only potential occupants, or “realizers,” of these causal roles.²⁸ To use a stock example, for an organism to be in pain is for it to be in some internal state that is typically caused by tissue damage and that typically causes groans, winces, and other characteristic pain behavior. In this sense being in pain is said to be a second-order property: for a system x to have this property is for x to have some first-order property P that satisfies a certain condition D , where in the present case D specifies that P has pain’s typical causes and typical effects.

More generally, we can explain the idea of a second-order property in the following way.²⁹ Let \mathbf{B} be a set of properties; these are our first-order (or “base”) properties. They are not first-order in any absolute sense; they may well be second-order relative to

another set of properties.³⁰ When mental properties are to be generated out of **B** as second-order properties, we must of course take **B** to consist of nonmental properties (including physico-chemical, biological, and behavioral properties³¹). We then have this:

F is a second-order property over set **B** of base (or first-order) properties iff *F* is the property of having some property *P* in **B** such that $D(P)$, where *D* specifies a condition on members of **B**.

Second-order properties therefore are second-order in that they are generated by *quantification*—existential quantification in the present case—over the base properties. We may call the base properties satisfying condition *D* the realizers of second-order property *F*. For example, if the base set **B** comprises colors, then *the property of having a primary color* can be thought of as a second-order property: having a property *P* in **B** such that $P = \text{red}$ or $P = \text{blue}$ or $P = \text{green}$.³² Thus being red, being blue, and being green are the three realizers of having a primary color. If **B** is a set of minerals,³³ being jade can be thought of as the second-order property of being a mineral that is pale green or white in color and fit for use as gemstones or for carving. This second-order property has two known realizers, jadeite and nephrite.

We need to say something about the vocabulary allowed for formulating condition *D*; for present purposes we will assume that the causal/nomological relation (holding for properties—or property instances, to be exact) is available, in addition to the usual logical expressions and appropriate descriptive terms (e.g., those referring to members of **B**). We may now explain *functional properties* over **B** as those second-order properties over **B** whose specification *D* involves the causal/nomic relation. That is, functional properties are second-order properties defined in terms of causal/nomic relations among first-order properties. An example of a functional property is dormitivity³⁴: a substance has this property just in case it has a chemical property that causes people to sleep. Both Valium and Seconal have dormitivity but in

virtue of different first-order (chemical) realizers—diazepam and secobarbital, respectively. Or consider *water-solubility*: something has this property just in case it has some property *P* such that when it is immersed in water *P* causes it to dissolve. This conception of functional property accords well with the standard usage in the functionalist literature. On the functionalist conception, mental properties are specified by causal roles, that is, in terms of causal relations holding for first-order physical properties (including biological and behavioral properties). In this sense mental properties turn out to be *extrinsic* or *relational* properties of individuals that have them. To be in a mental state is to be in a state with such-and-such as its typical causes and such-and-such as its typical effects. Whether or not a given property qualifies as an occupant of a specified role—that is, whether or not it is a realizer of a functional property—depends essentially on its causal/nomological relations to other properties, not on its intrinsic character. Intrinsic characters do matter of course, but only because of their capacity to get causally hooked up with other properties. Thus we may think of intrinsic characters as representing the causal potentials of their bearers (relative to the prevailing laws) which help determine whether or not their bearers instantiate a given functional property. On the other hand, networks of causal/nomological relations are constitutive of functional properties.³⁵

If mental properties are functional properties, they are therefore not tied, definitionally or constitutively, to the compositional/structural details of their realizers. Whether a given property realizes a given functional property is a contingent, empirical question; any base properties with the right causal/nomological relations to other properties can serve as its realizers. And any mechanism that gets activated by the right input and that, when activated, triggers the right response serves as a realizer of a psychological capacity or function. It has long been a platitude in philosophy of mind/psychology, and philosophical discussions of the special sciences, that mental properties and other special science properties can have extremely diverse realizers in different species and structures. This observation has

promoted a certain view about the nature of psychology and cognitive science, namely that the formal/abstract character of mental properties, standardly taken to be a consequence of their multiple realizability, is just what makes cognitive science possible—a scientific investigation of cognitive properties *as such*, across the diverse biological species and perhaps nonbiological cognitive systems, independently of the particulars of their physical implementation. In fact some have even speculated about the possibility of nonphysical realizations of psychologies; it is a seductive thought that there may be contingent empirical laws of cognition, or psychology, that are valid for *cognizers as such*, whether they are protein-based biological organisms like us and other earthly creatures, electromechanical robots, noncarbon-based intelligent extraterrestrials, immaterial Cartesian souls, heavenly angels, or even the omniscient one itself! (This surely takes the idea of “rational psychology” too far.) Even when we bring in the materialist constraint of physical realizationism, the idea of universal laws of cognition and psychology, applicable to all nomologically possible physical systems with cognitive capacities, is heady stuff, indeed.³⁶

Whether a given physical property P is a realizer of a mental property M depends on the nature of the system in which P is embedded,³⁷ since in psychology the input-output behavior of the total system is what is of concern, and the causal role that P plays will depend on the makeup (“causal wiring”) of the system as a whole. For example, whether or not tissue damage will cause the nociceptive neurons to fire in a given organism obviously depends on the organism’s neural organization, and whether or not the firing of these neural fibers will trigger appropriate escape behavior will again depend on the organism’s neural and motor systems. So the same property P , when embedded in a different system, may not realize M . Conversely, there may well be functional substitutes for P in the following sense: if for some reason the normal mechanism for instantiating P in an organism turns dysfunctional, another mechanism with appropriate causal capacities may be able to take its place and supply a near-enough realizer of M for that organism.

The status of P as a realizer of M varies along another dimension as well: since P 's credentials as M 's realizer depend on its causal/nomic relations to other properties, if laws of nature should vary, thereby altering P 's causal potential, that could affect P 's status as a realizer of M . P realizes M in this and other nomologically similar worlds; however, in worlds in which different laws hold at the level of M 's base domain, thereby generating different causal structures in those worlds, P may fail to satisfy the functional specification definitive of M . In such worlds, M may have realizers entirely different from its realizers in this world, or it may have no realizers at all.³⁸

Although the realization relation can shift in these ways, it is also important to note its constancy. Once the system's physical constitution and the prevailing laws of nature are fixed, that fixes whether or not P realizes M in that system. That is to say, if P realizes M in system s , then P will realize M in all systems which are subject to the same laws and which are relevantly indiscernible from s —that is, in respect of nomic properties. If, as most of us would accept, the microstructure of a system determines its causal/nomic properties, it follows that with laws held constant, the realization relation remains invariant for systems with similar microstructures.

Physical Realizationism Explains Mind-Body Supervenience

Consider a class S of systems sharing a relevantly similar microstructure. Biological conspecifics may constitute such a class. Suppose that P realizes M in systems of kind S . From the definition of realization, it follows that P is sufficient for M (that is, if a system of kind S instantiates P at t , it instantiates M at t)—in fact, given the nomological constancy just noted of the realization relation, it follows that P is nomologically sufficient for M . Thus, if $\langle P_1, \dots, P_n \rangle$ is a realization of $\langle M_1, \dots, M_n \rangle$, in the sense that each P_i is a realizer of M_i , it follows that the M s are supervenient on the P s. Physical realizationism therefore entails the supervenience thesis. Given the relativity of the realization relation to prevailing laws, the entailed supervenience

thesis has only the force of nomological necessity, not that of full metaphysical or logical/conceptual necessity. Thus physically indiscernible systems in worlds with different laws may not instantiate the same psychology.

This means that physical realizationism would give us an explanation of the supervenience thesis: the mental supervenes on the physical because mental properties are second-order functional properties with physical realizers (and no nonphysical realizers). And we have an explanation of mental-physical correlations. Why is it that whenever P is realized in a system s , it instantiates mental property M ? The answer is that by definition, having M is having a property with causal specification D , and in systems like s , P is the property (or one of the properties) meeting specification D . For systems like s , then, having M *consists in* having P . It isn't that when certain systems instantiate P , mental property M magically emerges or supervenes (in the dictionary sense of "supervene"). It is rather that having M for these systems, simply *is* having P . We might even say, using a familiar if shopworn reductive idiom, that having M , for these systems, is "nothing over and above" having P . Note that all of these explanations are nomological explanations—they depend on the fact that a certain set of laws prevail in our world, for these laws ultimately determine what physical properties are realizers of a given mental property. From a metaphysical point of view, therefore, the idea that mental properties are realized by physical properties goes considerably beyond such ideas as that mental properties have "physical correlates" or "neural substrates," that they have "physical supervenience bases," and the like; unlike "realization," these ideas are not capable of generating an explanation of why a given mental property arises out of, or correlates with, certain physical properties, and do not warrant reductive talk like "Having M , for appropriate systems, *consists in*, or *just is*, having P ."

These considerations, I believe, point to a conception of reduction that accords well with the paradigm of reduction in science. To reduce a property, or phenomenon, we first construe it—or reconstrue it—functionally, in terms of its causal/nomic rela-

tions to other properties and phenomena. To reduce temperature, we must first stop thinking of it as an intrinsic property but construe it as an extrinsic property characterized relationally, in terms of causal/nomic relations, perhaps something like this: it is that magnitude of an object that increases (or is caused to increase) when the object is in contact with another with a higher degree of it, that, when high, causes a ball of wax in the vicinity to melt, that causes the sensation of warmth or cold in humans, that, when extremely low, can make steel brittle, that, when extremely high, can turn steel into a molten state—you get the idea. Here is another example: the gene is that mechanism in a biological organism causally responsible for the transmission of heritable characteristics from parents to offsprings. To be transparent is to have the kind of molecular structure that causes light to pass through intact. And so on. We then find properties or mechanisms, often at the microlevel, that satisfy these causal/nomic specifications and thereby fill the specified causal roles. Multiple realization and nomic relativity obtain in these cases as well. Temperature may be one thing in gases but something else in solids, plasmas, and vacuums. The DNA molecule is the realizer of the gene, but in worlds in which different basic laws prevail molecules of another kind may perform the causal functions definitive of the gene. Reductions therefore are doubly relative: in systems with different structures, the underlying mechanisms realizing the reduced property may vary, and reductions remain valid only when the basic laws of nature are held constant—that is, only for nomologically possible worlds (relative to the reference world).³⁹

What has just been described differs in important ways from the standard model of theory reduction that has dominated the discussion of reduction, in particular, the possibility of mind-body reduction. This is Ernest Nagel's model of intertheoretic reduction whose principal focus is the derivation of laws.⁴⁰ According to Nagel, reduction is basically a proof procedure, consisting in the logical-mathematical derivation of the laws of the target theory from those of the base theory, taken in conjunction with "bridge laws" connecting the predicates of the two theories.

Standardly, these correlating bridge laws are taken to be biconditional in form, providing each property in the domain of the theory to be reduced with a nomologically coextensive property in the reduction base. For mind-body reduction Nagel's model requires that each mental property be provided with a nomologically coextensive physical property, across all species and structure types. This has made mind-body reductionism—in fact all reductionisms—an easy target. As everyone knows, the most influential antireductionist argument, one that had a decisive role in starting off the antireductionist bandwagon which is still going strong, is based on the claim that on account of the multiple realizability of mental properties, mental properties fail to have coextensions in the physical domain, thereby making mind-body bridge laws unavailable for Nagelian reduction. This argument was then generalized in defense of a general antireductionist position in regard to all special sciences.⁴¹ For three decades the battle over reductionisms has been fought on the question whether or not appropriate bridge laws are available for the domains involved.

But this is the wrong battlefield on which to contest the issue of reduction. What has gone largely unappreciated is the fact that the Nagel model of reduction is in effect the Hempelian D-N model of scientific explanation applied to intertheoretic contexts. Just as Hempelian explanation consists in the derivation of the statement describing the phenomenon to be explained from laws taken together with auxiliary premises describing relevant initial conditions, Nagelian reduction is accomplished in the derivation of the target theory from the base theory taken in conjunction with bridge laws as auxiliary premises. It is therefore more than a little surprising that while the D-N model of explanation has had few committed adherents for over three decades, Nagel's derivational model of reduction is still serving as the dominant standard in discussions of reduction and reductionism.⁴² I believe that Nagelian uniform reductions based on universal biconditional bridge laws are extremely rare (if any exists) in the sciences⁴³—especially, in the case of microreductions—and that the kind of model adumbrated above is not only more realistic

but also, as we will see in a later lecture, more appropriate from a metaphysical point of view. If this is right, the reducibility of a property critically depends on its *functionalizability*—whether or not it can be construed as a second-order functional property over properties in the base domain—not on the availability of bridge laws. Bridge laws are neither necessary nor sufficient for reduction.

We will later take up this and related questions, in connection with mind-body reduction and reductionism (chapter 4). Today I have tried to set the stage for more detailed discussion of some central issues that make up the mind-body problem, in particular the problems of mental causation and mind-body reduction. But several things have become clear: (1) if we are to understand the mind-body relation, we need a positive explanatory account of how mental properties are related to physical properties; (2) Davidson's anomalous monism fails to do this, and likewise with any other similar form of "token physicalism;" (3) even when mind-body supervenience is invoked, that does not help to generate such an account, for the supervenience itself is in need of an explanation; and (4) physical realizationism, whether or not it is ultimately correct, does promise such an account—at least, it has the right form and content. In particular, it can yield an explanation of mind-body supervenience and points to a more realistic model of reduction in terms of which the issue of mind-body reductionism can be more fruitfully debated.

References

-
- Alexander, Samuel. *Space, Time, and Deity*, vol. 2. London: Macmillan, 1920.
- Antony, Louise, M. "Anomalous Monism and the Problem of Explanatory Force." *Philosophical Review* 98 (1989): 153–187.
- Antony, Louise, M. "The Inadequacy of Anomalous Monism as a Realist Theory of Mind." In *Language, Mind, and Epistemology*, ed. G. Preyer, F. Siebelt, and A. Ulfig. Dordrecht: Kluwer, 1994.
- Antony, Louise M., and Joseph Levine. "Reduction with Autonomy." *Philosophical Perspectives* 11 (1997): 83–105.
- Armstrong, David. *A Materialist Theory of Mind*. New York: Humanities Press, 1968.
- Armstrong, David. *A Theory of Universals*, vol. 2. Cambridge: Cambridge University Press, 1978.
- Baker, Lynne Rudder. "Metaphysics and Mental Causation." In *Mental Causation*, ed. Heil and Mele.
- Baker, Lynne Rudder. *Explaining Attitudes*. Cambridge: Cambridge University Press, 1995.
- Beckermann, Ansgar, Hans Flohr, and Jaegwon Kim, eds. *Emergence or Reduction?* Berlin: De Gruyter, 1992.
- Bieri, Peter. "Trying out Epiphenomenalism." *Erkenntnis* 36 (1992): 283–309.
- Block, Ned. "Can the Mind Change the World?" In *Meaning and Method*, ed. George Boolos. Cambridge: Cambridge University Press, 1990.
- Block, Ned. "Introduction: What Is Functionalism?" In *Readings in Philosophy of Psychology*, vol. 1, ed. Block. Cambridge: Harvard University Press, 1980.
- Block, Ned. "Anti-Reductionism Slaps Back." *Philosophical Perspectives* 11 (1997): 107–132.
- Broad, C. D. *The Mind and Its Place in Nature*. London: Routledge and Kegan Paul, 1925.
- Burge, Tyler. "Mind-Body Causation and Explanatory Practice." In *Mental Causation*, ed. John Heil and Alfred Mele.
- Burge, Tyler. "Individualism and the Mental." *Midwest Studies in Philosophy* 4 (1979): 73–121.
- Burge, Tyler. "Philosophy of Language and Mind: 1950–1990." *Philosophical Review* 101 (1992): 3–51.
- Causey, Robert. *Unity of Science*. Dordrecht: Reidel, 1977.

- Chalmers, David. *The Conscious Mind*. Oxford: Oxford University Press, 1996.
- Davidson, Donald. "Mental Events." 1970. Reprinted in Davidson, *Essays on Actions and Events*.
- Davidson, Donald. *Essays on Actions and Events*. Oxford: Oxford University Press, 1980.
- Davidson, Donald. "Actions, Reasons, and Causes." *Journal of Philosophy* 60 (1963). Reprinted in *Essays on Actions and Events*.
- Davidson, Donald. "The Individuation of Events." Reprinted in *Essays on Actions and Events*.
- Davidson, Donald. "Psychology as Philosophy." Reprinted in *Essays on Actions and Events*.
- Davidson, Donald. "Thinking Causes." In *Mental Causation*, ed. John Heil and Alfred Mele.
- Descartes, René. *The Philosophical Writings of Descartes*, vol. 2, ed. John Cottingham, Robert Stoothoff, and Dugald Murdoch. Cambridge: Cambridge University Press, 1985.
- Dray, William H. *Laws and Explanations in History*. Oxford: Oxford University Press, 1957.
- Dretske, Fred. *Naturalizing the Mind*. Cambridge: MIT Press, 1995.
- Endicott, Ronald. "Constructural Plasticity." *Philosophical Studies* 74 (1994): 51–75.
- Feigl, Herbert. "The 'Mental' and the 'Physical'." *Minnesota Studies in the Philosophy of Science*, vol. 2, ed. Herbert Feigl, Grover Maxwell, and Michael Scriven. Minneapolis: University of Minnesota Press, 1958.
- Fodor, Jerry A. "Special Sciences, or the Disunity of Science as a Working Hypothesis." *Synthese* 28 (1974): 97–115. Reprinted in *Representations*.
- Fodor, Jerry A. "Making Mind Matter More." *Philosophical Topics* 17 (1989): 59–79. Reprinted in *A Theory of Content and Other Essays*. Cambridge: MIT Press, 1990.
- Fodor, Jerry A. *Representations*. Cambridge: MIT Press, 1981.
- Ginet, Ginet. *On Action*. Cambridge: Cambridge University Press, 1990.
- Goldberg, Bruce. "The Correspondence Hypothesis." *Philosophical Review* 77 (1968): 439–454.
- Heil, John. *The Nature of True Minds*. Cambridge: Cambridge University Press, 1992.
- Heil, John, and Alfred Mele, eds. *Mental Causation*. Oxford: Clarendon, 1993.
- Hempel, Carl G., and Paul Oppenheim. "Studies in the Logic of Explanation." Reprinted in Hempel, *Aspects of Scientific Explanation*. New York: Free Press, 1965.
- Honderich, Ted. "The Argument for Anomalous Monism." *Analysis* 42 (1982): 59–64.
- Horgan, Terence. "From Supervenience to Superdupervenience." *Mind* 102 (1993): 555–586.
- Horgan, Terence. "Supervenience and Cosmic Hermeneutics." *Southern Journal of Philosophy* 22 (1984), suppl: 19–38.

- Horgan, Terence. "Mental Quausation." *Philosophical Perspectives* 3 (1989): 47–76.
- Horgan, Terence. "Supervenient Qualia." *Philosophical Review* 96 (1987): 491–520.
- Horgan, Terence. "Kim on Mental Causation and Causal Exclusion." *Philosophical Perspectives* 11 (1997): 165–184.
- Horgan, Terence, and Mark Timmons. "Troubles on Moral Twin Earth: Moral Queerness Revisited." *Synthese* 92 (1992): 221–260.
- Jackson, Frank, and Philip Pettit. "Functionalism and Broad Content." *Mind* 97 (1988): 381–400.
- Jackson, Frank, and Philip Pettit. "Program Explanation: A General Perspective." *Analysis* 50 (1990): 107–117.
- Jacob, Pierre. *What Minds Can Do*. Cambridge: Cambridge University Press, 1997.
- Kim, Jaegwon. "Does the Problem of Mental Causation Generalize?" *Proceedings of the Aristotelian Society* 97 (1997): 281–297.
- Kim, Jaegwon. "Psychophysical Laws." Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. "What is the Problem of Mental Causation?" *Norms and Structures in Science*, ed. M. L. Dalla Chiara et al. (Dordrecht: Kluwer, 1997).
- Kim, Jaegwon. "Concepts of Supervenience." Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. "Mental Causation: What? Me Worry?" *Philosophical Issues* 6 (1995): 123–151.
- Kim, Jaegwon. "Epiphenomenal and Supervenient Causation." *Midwest Studies in Philosophy* 9 (1984): 257–270. Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. "Mechanism, Purpose, and Explanatory Exclusion." *Philosophical Perspectives* 3 (1989): 77–108. Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. "Multiple Realization and the Metaphysics of Reduction." Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. "Psychophysical Supervenience." *Philosophical Studies* 41 (1982): 51–70. Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. "Self-Understanding and Rationalizing Explanations." *Philosophia Naturalis* 82 (1984): 309–20.
- Kim, Jaegwon. "The Myth of Nonreductive Physicalism." *Proceedings and Addresses of the American Philosophical Association* 63 (1989): 31–47. Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. "The Nonreductivist's Troubles with Mental Causation." In *Supervenience and Mind*.
- Kim, Jaegwon. "'Downward Causation' in Emergentism and Nonreductive Physicalism." In *Emergence or Reduction?*, ed. A. Beckermann, H. Flohr, and J. Kim.
- Kim, Jaegwon. "'Strong' and 'Global' Supervenience Revisited." Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. "Supervenience as a Philosophical Concept." Reprinted in *Supervenience and Mind*.
- Kim, Jaegwon. *Supervenience and Mind*. Cambridge: Cambridge University Press, 1993.

- Kim, Jaegwon. *Philosophy of Mind*. Boulder, CO: Westview, 1996.
- Kim, Jaegwon. "The Mind-Body Problem after 50 Years." Delivered at the Royal Institute of Philosophy, October, 1996.
- Kim, Jaegwon. "The Mind-Body Problem: Taking Stock after 40 Years," *Philosophical Perspectives*, 11 (1997): 185–207.
- Kim, Jaegwon. "Making Sense of Emergence." Presented at the 1997 Oberlin Colloquium in Philosophy (forthcoming in *Philosophical Studies*).
- Kripke, Saul. *Naming and Necessity*. Cambridge: Harvard University Press, 1980.
- LePore, Ernest, and Barry Loewer. "Mind Matters." *Journal of Philosophy* 93 (1987): 630–642.
- Levine, Joseph. "On Leaving out What It's Like." In *Consciousness*, ed. Martin Davies and Glyn W. Humphreys. Oxford: Blackwell, 1993.
- Lewis, David. "Counterfactual Dependence and Time's Arrow." Reprinted in *Philosophical Papers II*. Oxford: Oxford University Press, 1986.
- Lewis, David. "Causal Explanation." In *Philosophical Papers II*. Oxford: Oxford University Press, 1986.
- Lewis, David. "Causation." Reprinted in *Philosophical Papers II*.
- Louch, A. R. *Explanation and Human Action*. Oxford: Blackwell, 1968.
- Lycan, William G. *Consciousness*. Cambridge: MIT Press, 1987.
- Malcolm, Norman. *Memory and Mind*. Ithaca: Cornell University Press, 1977.
- Marr, David. *Vision*. New York: Freeman Press, 1982.
- McLaughlin, Brian. "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical." *Philosophical Perspectives* 3 (1989): 109–135.
- McLaughlin, Brian. "Varieties of Supervenience." In *Supervenience: New Essays*, ed. E. Savellos and Umit D. Yalcin. Cambridge: Cambridge University Press, 1995.
- McLaughlin, Brian. "The Rise and Fall of British Emergentism." In *Emergence or Reduction?*, ed. Beckerman, Flohr, and Kim.
- Melden, A. I. *Free Action*. London: Routledge and Kegan Paul, 1961.
- Morgan, C. Lloyd. *Emergent Evolution*. London: Williams and Norgate, 1923.
- Nagel, Ernest. *The Structure of Science*. New York: Harcourt, Brace, 1961.
- Oppenheim, Paul, and Hilary Putnam. "Unity of Science as a Working Hypothesis." *Minnesota Studies in the Philosophy of Science*, vol. 2, ed. H. Feigl, G. Maxwell, and M. Scriven. Minneapolis: University of Minnesota Press, 1958.
- Place, U. T. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47, Part I (1956): 44–50.
- Post, John. *The Faces of Existence*. Ithaca: Cornell University Press, 1987.
- Prior, Elizabeth, Robert Pargetter, and Frank Jackson. "Three Theses About Dispositions." *American Philosophical Quarterly* 19 (1982): 251–257.
- Putnam, Hilary. "The Meaning of 'Meaning'." In *Philosophical Papers*, vol. 2. Cambridge: Cambridge University Press, 1975.
- Putnam, Hilary. "On Properties." In *Philosophical Papers*, vol. 1. Cambridge: Cambridge University Press, 1975.

- Putnam, Hilary. "Psychological Predicates." First published in 1968 and later Reprinted under the title "The Nature of Mental States." In Putnam, *Collected Papers II*. Cambridge: Cambridge University Press, 1975.
- Putnam, Hilary. "Minds and Machines." In *Dimensions of Mind*, ed. Sydney Hook. New York: New York University Press, 1960.
- Putnam, Hilary. *Representation and Reality*. Cambridge: MIT Press, 1988.
- Ryle, Gilbert. *The Concept of Mind*. London: Hutchinson and Company, Ltd., 1949.
- Salmon, Wesley C. *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press, 1984.
- Savellos, Elias, and Umit Yalcin, eds. *Supervenience: New Essays*. Cambridge: Cambridge University Press, 1995.
- Searle, John R. *The Rediscovery of the Mind*. Cambridge: MIT Press, 1992.
- Searle, John R. "Consciousness, the Brain and the Connection Principle: A Reply." *Philosophy and Phenomenological Research* 55 (1995): 217–232.
- Shoemaker, Sydney. "Some Varieties of Functionalism." Reprinted in *Identity, Cause, and Mind*. Cambridge: Cambridge University Press, 1984.
- Shoemaker, Sydney. "Causality and Properties." Reprinted in *Identity, Cause, and Mind*. Cambridge: Cambridge University Press, 1984.
- Smart, J. J. C. "Sensations and Brain Processes." *Philosophical Review* 68 (1959): 141–156.
- Sosa, Ernest. "Mind-Body Interaction and Supervenient Causation." *Midwest Studies in Philosophy* 9 (1984): 271–81.
- Stich, Stephen P. *From Folk Psychology to Cognitive Science*. Cambridge: MIT Press, 1983.
- Stich, Stephen P. "Autonomous Psychology and the Belief-Desire Thesis." *The Monist* 61 (1978): 573–591.
- Stoutland, Frederick. "Oblique Causation and Reasons for Action." *Synthese* 43 (1980): 351–67.
- Tye, Michael. *Ten Problems of Consciousness*. Cambridge: MIT Press, 1995.
- Van Gulick, Robert. "Three Bad Arguments for Intentional Property Epiphenomenalism." *Erkenntnis* 36 (1992).
- Van Gulick, Robert. "Nonreductive Materialism and the Nature of Intertheoretic Constraint." In *Emergence or Reduction*, ed. Beckermann, Flohr, and Kim.
- Varela, Francisco, Evan Thompson, and Eleanor Rosch. *The Embodied Mind*. Cambridge: MIT Press, 1993.
- von Wright, G. H. *Explanation and Understanding*. Ithaca: Cornell University Press, 1971.
- Wilson, George. *The Intentionality of Human Action*. Stanford: Stanford University Press, 1989.

