Philosophy

Robert A. Wilson

The areas of philosophy that contribute to and draw on the cognitive sciences are various; they include the philosophy of mind, science, and language; formal and philosophical logic; and traditional metaphysics and epistemology. The most direct connections hold between the philosophy of mind and the cognitive sciences, and it is with classical issues in the philosophy of mind that I begin this introduction (section 1). I then briefly chart the move from the rise of materialism as the dominant response to one of these classic issues, the mind-body problem, to the idea of a science of the mind. I do so by discussing the early attempts by introspectionists and behaviorists to study the mind (section 2). Here I focus on several problems with a philosophical flavor that arise for these views, problems that continue to lurk backstage in the theater of contemporary cognitive science.

Between these early attempts at a science of the mind and today's efforts lie two general, influential philosophical traditions, ordinary language philosophy and logical positivism. In order to bring out, by contrast, what is distinctive about the contemporary naturalism integral to philosophical contributions to the cognitive sciences, I sketch the approach to the mind in these traditions (section 3). And before getting to contemporary naturalism itself I take a quick look at the philosophy of science, in light of the legacy of positivism (section 4).

In sections 5 through 7 I get, at last, to the mind in cognitive science proper. Section 5 discusses the conceptions of mind that have dominated the contemporary cognitive sciences, particularly that which forms part of what is sometimes called "classic" cognitive science and that of its connectionist rival. Sections 6 and 7 explore two specific clusters of topics that have been the focus of philosophical discussion of the mind over the last 20 years or so, folk psychology and mental content. The final sections gesture briefly at the interplay between the cognitive sciences and logic (section 8) and biology (section 9).

1 Three Classic Philosophical Issues About the Mind

i. The Mental-Physical Relation

The relation between the mental and the physical is the deepest and most recurrent classic philosophical topic in the philosophy of mind, one very much alive today. In due course, we will come to see why this topic is so persistent and pervasive in thinking about the mind. But to convey something of the topic's historical significance let us begin with a classic expression of the puzzling nature of the relation between the mental and the physical, the MIND-BODY PROBLEM.

This problem is most famously associated with RENÉ DESCARTES, the preeminent figure of philosophy and science in the first half of the seventeenth century. Descartes combined a thorough-going mechanistic theory of nature with a *dualistic* theory of the nature of human beings that is still, in general terms, the most widespread view held by ordinary people outside the hallowed halls of academia. Although nature, including that of the human body, is material and thus completely governed by basic principles of mechanics, human beings are special in that they are composed both of material and nonmaterial or mental stuff, and so are not so governed. In Descartes's own terms, people are essentially a combination of mental substances (minds) and material substances (bodies). This is Descartes's *dualism*. To put it in more commonsense terms, people have both a mind and a body.

Although dualism is often presented as a possible solution to the mind-body problem, a possible position that one might adopt in explaining how the mental and physical are related, it serves better as a way to bring out why there is a "problem" here at all. For if the mind is one type of thing, and the body is another, how do these two

types of things interact? To put it differently, if the mind really is a nonmaterial substance, lacking physical properties such as spatial location and shape, how can it be both the cause of effects in the material world—like making bodies move—and itself be causally affected by that world—as when a thumb slammed with a hammer (bodily cause) causes one to feel pain (mental effect)? This problem of causation between mind and body has been thought to pose a largely unanswered problem for Cartesian dualism.

It would be a mistake, however, to assume that the mind-body problem in its most general form is simply a consequence of dualism. For the general question as to how the mental is related to the physical arises squarely for those convinced that some version of materialism or PHYSICALISM must be true of the mind. In fact, in the next section, I will suggest that one reason for the resilience and relevance of the mind-body problem has been the *rise* of materialism over the last fifty years.

Materialists hold that all that exists is material or physical in nature. Minds, then, are somehow or other composed of arrangements of physical stuff. There have been various ways in which the "somehow or other" has been cashed out by physicalists, but even the view that has come closest to being a consensus view among contemporary materialists—that the mind *supervenes* on the body—remains problematic. Even once one adopts materialism, the task of articulating the relationship between the mental and the physical remains, because even physical minds have special properties, like intentionality and consciousness, that require further explanation. Simply proclaiming that the mind is not made out of distinctly mental substance, but is material like the rest of the world, does little to explain the features of the mind that seem to be distinctively if not uniquely features of physical minds.

ii. The Structure of the Mind and Knowledge

Another historically important cluster of topics in the philosophy of mind concerns what is in a mind. What, if anything, is distinctive of the mind, and how is the mind structured? Here I focus on two dimensions to this issue.

One dimension stems from the RATIONALISM VS. EMPIRICISM debate that reached a high point in the seventeenth and eighteenth centuries. Rationalism and empiricism are views of the nature of human knowledge. Broadly speaking, empiricists hold that all of our knowledge derives from our sensory, experiential, or empirical interaction with the world. Rationalists, by contrast, hold the negation of this, that there is some knowledge that does not derive from experience.

Since at least our paradigms of knowledge—of our immediate environments, of common physical objects, of scientific kinds—seem obviously to be based on sense experience, empiricism has significant intuitive appeal. Rationalism, by contrast, seems to require further motivation: minimally, a list of knowables that represent a prima facie challenge to the empiricist's global claim about the foundations of knowledge. Classic rationalists, such as Descartes, Leibniz, Spinoza, and perhaps more contentiously KANT, included knowledge of God, substance, and abstract ideas (such as that of a triangle, as opposed to ideas of particular triangles). Empiricists over the last three hundred years or so have either claimed that there was nothing to know in such cases, or sought to provide the corresponding empiricist account of how we could know such things from experience.

The different views of the sources of knowledge held by rationalists and empiricists have been accompanied by correspondingly different views of the mind, and it is not hard to see why. If one is an empiricist and so holds, roughly, that there is nothing in the mind that is not first in the senses, then there is a fairly literal sense in which *ideas,* found in the mind, are complexes that derive from *impressions* in the senses. This in turn suggests that the processes that constitute cognition are themselves elaborations of those that constitute perception, that is, that cognition and perception differ only in degree, not kind. The most commonly postulated mechanisms governing these processes are *association* and *similarity,* from Hume's laws of association to feature-extraction in contemporary connectionist networks. Thus, the mind tends to be viewed by empiricists as a *domain-general* device, in that the principles that govern its opera-

tion are constant across various types and levels of cognition, with the common empirical basis for all knowledge providing the basis for parsimony here.

By contrast, in denying that all knowledge derives from the senses, rationalists are faced with the question of what other sources there are for knowledge. The most natural candidate is the mind itself, and for this reason rationalism goes hand in hand with NATIVISM about both the source of human knowledge and the structure of the human mind. If some ideas are innate (and so do not need to be derived from experience), then it follows that the mind already has a relatively rich, inherent structure, one that in turn limits the malleability of the mind in light of experience. As mentioned, classic rationalists made the claim that certain ideas or CONCEPTS were innate, a claim occasionally made by contemporary nativists—most notably Jerry Fodor (1975) in his claim that *all* concepts are innate. However, contemporary nativism is more often expressed as the view that certain implicit knowledge that we have or principles that govern how the mind works—most notoriously, linguistic knowledge and principles—are innate, and so not learned. And because the types of knowledge that one can have may be endlessly heterogeneous, rationalists tend to view the mind as a *domain-specific* device, as one made up of systems whose governing principles are very different. It should thus be no surprise that the historical debate between rationalists and empiricists has been revisited in contemporary discussions of the INNATENESS OF LANGUAGE, the MODULARITY OF MIND, and CONNECTIONISM.

A second dimension to the issue of the structure of the mind concerns the place of CONSCIOUSNESS among mental phenomena. From WILLIAM JAMES's influential analysis of the phenomenology of the stream of consciousness in his *The Principles of Psychology* (1890) to the renaissance that consciousness has experienced in the last ten years (if publication frenzies are anything to go by), consciousness has been thought to be the most puzzling of mental phenomena. There is now almost universal agreement that conscious mental states are a part of the mind. But how large and how important a part? Consciousness has sometimes been thought to exhaust the mental, a view often attributed to Descartes. The idea here is that everything mental is, in some sense, conscious or available to consciousness. (A version of the latter of these ideas has been recently expressed in John Searle's [1992: 156] *connection principle*: "all unconscious intentional states are in principle accessible to consciousness.")

There are two challenges to the view that everything mental is conscious or even available to consciousness. The first is posed by the *unconscious*. SIGMUND FREUD's extension of our common-sense attributions of belief and desire, our folk psychology, to the realm of the unconscious played and continues to play a central role in PSYCHOANALYSIS. The second arises from the conception of cognition as information processing that has been and remains focal in contemporary cognitive science, because such information processing is mostly *not* available to consciousness. If cognition so conceived is mental, then most mental processing is not available to consciousness.

iii. The First- and Third-Person Perspectives

Occupying center stage with the mind-body problem in traditional philosophy of mind is the *problem of other minds*, a problem that, unlike the mind-body problem, has all but disappeared from philosophical contributions to the cognitive sciences. The problem is often stated in terms of a contrast between the relatively secure way in which I "directly" know about the existence of *my own* mental states, and the far more epistemically risky way in which I must infer the existence of the mental states of others. Thus, although I can know about my own mental states simply by introspection and self-directed reflection, because this way of finding out about mental states is peculiarly first-person, I need some other type of evidence to draw conclusions about the mental states of others. Naturally, an agent's behavior is a guide to what mental states he or she is in, but there seems to be an epistemic gap between this sort of evidence and the attribution of the corresponding mental states that does not exist in the case of self-ascription. Thus the problem of other minds is chiefly an *epistemological* problem, sometimes expressed as a form of skepticism about the justification that we have for attributing mental states to others.

There are two reasons for the waning attention to the problem of other minds *qua problem* that derive from recent philosophical thought sensitive to empirical work in the cognitive sciences. First, research on introspection and SELF-KNOWLEDGE has raised questions about how "direct" our knowledge of our own mental states and of the SELF is, and so called into question traditional conceptions of first-person knowledge of mentality. Second, explorations of the THEORY OF MIND, ANIMAL COMMUNICATION, and SOCIAL PLAY BEHAVIOR have begun to examine and assess the sorts of attribution of mental states that are actually justified in empirical studies, suggesting that third-person knowledge of mental states is not as limited as has been thought. Considered together, this research hints that the contrast between first- and third-person knowledge of the mental is not as stark as the problem of other minds seems to intimate.

Still, there is something distinctive about the first-person perspective, and it is in part as an acknowledgment of this, to return to an earlier point, that consciousness has become a hot topic in the cognitive sciences of the 1990s. For whatever else we say about consciousness, it seems tied ineliminably to the first-person perspective. It is a state or condition that has an irreducibly *subjective* component, something with an essence to be experienced, and which presupposes the existence of a subject of that experience. Whether this implies that there are QUALIA that resist complete characterization in materialist terms, or other limitations to a science of the mind, remain questions of debate.

See also ANIMAL COMMUNICATION; CONCEPTS; CONNECTIONISM, PHILOSOPHICAL ISSUES; CONSCIOUSNESS; CONSCIOUSNESS, NEUROBIOLOGY OF; DESCARTES, RENÉ; FREUD, SIGMUND; INNATENESS OF LANGUAGE; JAMES, WILLIAM; KANT, IMMANUEL; MIND-BODY PROBLEM; MODULARITY OF MIND; NATIVISM; NATIVISM, HISTORY OF; PHYSICALISM; PSYCHOANALYSIS, CONTEMPORARY VIEWS; PSYCHOANALYSIS, HISTORY OF; QUALIA; RATIONALISM VS. EMPIRICISM; SELF; SELF-KNOWLEDGE; SOCIAL PLAY BEHAVIOR; THEORY OF MIND

2 *From Materialism to Mental Science*

In raising issue *i.*, the mental-physical relation, in the previous section, I implied that materialism was the dominant ontological view of the mind in contemporary philosophy of mind. I also suggested that, if anything, general convergence on this issue has intensified interest in the mind-body problem. For example, consider the large and lively debate over whether contemporary forms of materialism are compatible with genuine MENTAL CAUSATION, or, alternatively, whether they commit one to EPIPHENOMENALISM about the mental (Kim 1993; Heil and Mele 1993; Yablo 1992). Likewise, consider the fact that despite the dominance of materialism, some philosophers maintain that there remains an EXPLANATORY GAP between mental phenomena such as consciousness and any physical story that we are likely to get about the workings of the brain (Levine 1983; cf. Chalmers 1996). Both of these issues, very much alive in contemporary philosophy of mind and cognitive science, concern the mind-body problem, even if they are not always identified in such old-fashioned terms.

I also noted that a healthy interest in the first-person perspective persists within this general materialist framework. By taking a quick look at the two major initial attempts to develop a systematic, scientific understanding of the mind—late nineteenth-century introspectionism and early twentieth-century behaviorism—I want to elaborate on these two points and bring them together.

Introspectionism was widely held to fall prey to a problem known as the *problem of the homunculus*. Here I argue that behaviorism, too, is subject to a variation on this very problem, and that both versions of this problem continue to nag at contemporary sciences of the mind.

Students of the history of psychology are familiar with the claim that the roots of contemporary psychology can be dated from 1879, with the founding of the first experimental laboratory devoted to psychology by WILHELM WUNDT in Leipzig, Germany. As an *experimental* laboratory, Wundt's laboratory relied on the techniques introduced and refined in physiology and psychophysics over the preceding fifty years

by HELMHOLTZ, Weber, and Fechner that paid particular attention to the report of SENSATIONS. What distinguished Wundt's as a laboratory of *psychology* was his focus on the data reported in consciousness via the first-person perspective; psychology was to be the science of immediate experience and its most basic constituents. Yet we should remind ourselves of how restricted this conception of psychology was, particularly relative to contemporary views of the subject.

First, Wundt distinguished between mere INTROSPECTION, first-person reports of the sort that could arise in the everyday course of events, and experimentally manipulable self-observation of the sort that could only be triggered in an experimental context. Although Wundt is often thought of as the founder of an introspectionist methodology that led to a promiscuous psychological ontology, in disallowing mere introspection as an appropriate method for a science of the mind he shared at least the sort of restrictive conception of psychology with *both* his physiological predecessors and his later behaviorist critics.

Second, Wundt thought that the vast majority of ordinary thought and cognition was *not* amenable to acceptable first-person analysis, and so lay beyond the reach of a scientific psychology. Wundt thought, for example, that belief, language, personality, and SOCIAL COGNITION could be studied systematically only by detailing the cultural mores, art, and religion of whole societies (hence his four-volume *Völkerpsychologie* of 1900–1909). These studies belonged to the humanities (*Geisteswissenshaften*) rather than the experimental sciences (*Naturwissenschaften*), and were undertaken by anthropologists inspired by Wundt, such as BRONISLAW MALINOWSKI.

Wundt himself took one of his early contributions to be a solution of the mind-body problem, for that is what the data derived from the application of the experimental method to distinctly psychological phenomena gave one: correlations between the mental and the physical that indicated how the two were systematically related. The discovery of psychophysical laws of this sort showed how the mental was related to the physical. Yet with the expansion of the domain of the mental amenable to experimental investigation over the last 150 years, the mind-body problem has taken on a more acute form: just how do we get all that mind-dust from merely material mechanics? And it is here that the problem of the homunculus arises for introspectionist psychology after Wundt.

The problem, put in modern guise, is this. Suppose that one introspects, say, in order to determine the location of a certain feature (a cabin, for example) on a map that one has attempted to memorize (Kosslyn 1980). Such introspection is typically reported in terms of exploring a mental image with one's *mind's eye*. Yet we hardly want our psychological story to end there, because it posits a process (introspection) and a processor (the mind's eye) that themselves cry out for further explanation. The problem of the homunculus is the problem of leaving undischarged homunculi ("little men" or their equivalents) in one's *explanantia,* and it persists as we consider an elaboration on our initial introspective report. For example, one might well report forming a mental image of the map, and then scanning around the various features of the map, zooming in on them to discern more clearly what they are to see if any of them is the sought-after cabin. To take this introspective report seriously as a guide to the underlying psychological mechanisms would be to posit, minimally, an *imager* (to form the initial image), a *scanner* (to guide your mind's eye around the image), and a *zoomer* (to adjust the relative sizes of the features on the map). But here again we face the problem of the homunculus, because such "mechanisms" themselves require further psychological decomposition.

To be faced with the problem of the homunculus, of course, is not the same as to succumb to it. We might distinguish two understandings of just what the "problem" is here. First, the problem of the homunculus could be viewed as a problem specifically for introspectionist views of psychology, a problem that was never successfully met and that was principally responsible for the abandonment of introspectionism. As such, the problem motivated BEHAVIORISM in psychology. Second, the problem of the homunculus might simply be thought of as a challenge that *any* view that posits internal mental states must respond to: to show how to discharge all of the homunculi introduced in a way that is acceptably materialistic. So construed, the problem

remains one that has been with us more recently, in disputes over the psychological reality of various forms of GENERATIVE GRAMMAR (e.g., Stabler 1983); in the nativism that has been extremely influential in post-Piagetian accounts of COGNITIVE DEVELOPMENT (Spelke 1990; cf. Elman et al. 1996); and in debates over the significance of MENTAL ROTATION and the nature of IMAGERY (Kosslyn 1994; cf. Pylyshyn 1984: ch.8).

With Wundt's own restrictive conception of psychology and the problem of the homunculus in mind, it is with some irony that we can view the rise and fall of behaviorism as the dominant paradigm for psychology subsequent to the introspectionism that Wundt founded. For here was a view so deeply indebted to materialism and the imperative to explore psychological claims only by reference to what was acceptably experimental that, in effect, in its purest form it appeared to do away with the distinctively mental altogether! That is, because objectively observable behavioral responses to objectively measurable stimuli are all that could be rigorously explored, experimental psychological investigations would need to be significantly curtailed, relative to those of introspectionists such as Wundt and Titchener. As J. B. Watson said in his early, influential "Psychology as the Behaviorist Views It" in 1913, "Psychology as behavior will, after all, have to neglect but few of the really essential problems with which psychology as an introspective science now concerns itself. In all probability even this residue of problems may be phrased in such a way that refined methods in behavior (which certainly must come) will lead to their solution" (p. 177).

Behaviorism brought with it not simply a global conception of psychology but specific methodologies, such as CONDITIONING, and a focus on phenomena, such as that of LEARNING, that have been explored in depth since the rise of behaviorism. Rather than concentrate on these sorts of contribution to the interdisciplinary sciences of the mind that behaviorists have made, I want to focus on the central problem that faced behaviorism as a research program for reshaping psychology.

One of the common points shared by behaviorists in their philosophical and psychological guises was a commitment to an *operational* view of psychological concepts and thus a suspicion of any reliance on concepts that could not be operationally characterized. Construed as a view of scientific *definition* (as it was by philosophers), operationalism is the view that scientific terms must be defined in terms of observable and measurable operations that one can perform. Thus, an operational definition of "length," as applied to ordinary objects, might be: "the measure we obtain by laying a standard measuring rod or rods along the body of the object." Construed as a view of scientific *methodology* (as it was by psychologists), operationalism claims that the subject matter of the sciences should be objectively observable and measurable, by itself a view without much content.

The real bite of the insistence on operational definitions and methodology for psychology came via the application of operationalism to unobservables, for the various feelings, sensations, and other internal states reported by introspection, themselves unobservable, proved difficult to operationalize adequately. Notoriously, the introspective reports from various psychological laboratories produced different listings of the basic feelings and sensations that made up consciousness, and the lack of agreement here generated skepticism about the reliability of introspection as a method for revealing the structure of the mind. In psychology, this led to a focus on behavior, rather than consciousness, and to its exploration through observable stimulus and response: hence, behaviorism. But I want to suggest that this reliance on operationalism itself created a version of the problem of the homunculus for behaviorism. This point can be made in two ways, each of which offers a reinterpretation of a standard criticism of behaviorism. The first of these criticisms is usually called "philosophical behaviorism," the attempt to provide conceptual analyses of mental state terms exclusively in terms of behavior; the second is "psychological behaviorism," the research program of studying objective and observable behavior, rather than subjective and unobservable inner mental episodes.

First, as Geach (1957: chap. 4) pointed out with respect to belief, behaviorist analyses of individual folk psychological states are bound to fail, because it is only in concert with many other propositional attitudes that any given such attitude has

behavioral effects. Thus, to take a simple example, we might characterize the belief that it is raining as the tendency to utter "yes" when asked, "Do you believe that it is raining?" But one reason this would be inadequate is that one will engage in this verbal behavior only if one *wants* to answer truthfully, and only if one *hears* and *understands* the question asked, where each of the italicized terms above refers to some other mental state. Because the problem recurs in *every* putative analysis, this implies that a behavioristically acceptable construal of folk psychology is not possible. This point would seem to generalize beyond folk psychology to representational psychology more generally.

So, in explicitly attempting to do without internal mental representations, behaviorists themselves are left with mental states that must simply be assumed. Here we are not far from those undischarged homunculi that were the bane of introspectionists, especially once we recognize that the metaphorical talk of "homunculi" refers precisely to internal mental states and processes that themselves are not further explained.

Second, as Chomsky (1959: esp. p. 54) emphasized in his review of Skinner's *Verbal Behavior,* systematic attempts to operationalize psychological language invariably smuggle in a reference to the very mental processes they are trying to do without. At the most general level, the behavior of interest to the linguist, Skinner's "verbal behavior," is difficult to characterize adequately without at least an implicit reference to the sorts of psychological mechanism that generate it. For example, linguists are not interested in mere noises that have the same physical properties—"harbor" may be pronounced so that its first syllable has the same acoustic properties as an exasperated grunt—but in parts of speech that are taxonomized at least partially in terms of the surrounding mental economy of the speaker or listener.

The same seems true for *all* of the processes introduced by behaviorists—for example, stimulus control, reinforcement, conditioning—insofar as they are used to characterize complex, human behavior that has a natural psychological description (making a decision, reasoning, conducting a conversation, issuing a threat). What marks off their instances as behaviors *of the same kind* is not exclusively their physical or behavioral similarity, but, in part, the common, internal psychological processes that generate them, and that they in turn generate. Hence, the irony: behaviorists, themselves motivated by the idea of reforming psychology so as to generalize about objective, observable behavior and so avoid the problem of the homunculus, are faced with undischarged homunculi, that is, irreducibly mental processes, in their very own alternative to introspectionism.

The two versions of the problem of the homunculus are still with us as a Scylla and Charybdis for contemporary cognitive scientists to steer between. On the one hand, theorists need to avoid building the very cognitive abilities that they wish to explain into the models and theories they construct. On the other, in attempting to side-step this problem they also run the risk of masking the ways in which their "objective" taxonomic categories presuppose further internal psychological description of precisely the sort that gives rise to the problem of the homunculus in the first place.

See also BEHAVIORISM; COGNITIVE DEVELOPMENT; CONDITIONING; EPIPHENOMENALISM; EXPLANATORY GAP; GENERATIVE GRAMMAR; HELMHOLTZ, HERMANN; IMAGERY; INTROSPECTION; LEARNING; MALINOWSKI, BRONISLAW; MENTAL CAUSATION; MENTAL ROTATION; SENSATIONS; SOCIAL COGNITION; SOCIAL COGNITION IN ANIMALS; WUNDT, WILHELM

3 *A Detour Before the Naturalistic Turn*

Given the state of philosophy and psychology in the early 1950s, it is surprising that within twenty-five years there would be a thriving and well-focused interdisciplinary unit of study, cognitive science, to which the two are central. As we have seen, psychology was dominated by behaviorist approaches that were largely skeptical of positing internal mental states as part of a serious, scientific psychology. And Anglo-American philosophy featured two distinct trends, each of which made philosophy more insular with respect to other disciplines, and each of which served to reinforce the behaviorist orientation of psychology.

First, ordinary language philosophy, particularly in Great Britain under the influence of Ludwig Wittgenstein and J. L. Austin, demarcated distinctly philosophical problems as soluble (or dissoluble) chiefly by reference to what one would ordinarily say, and tended to see philosophical views of the past and present as the result of confusions in how philosophers and others come to use words that generally have a clear sense in their ordinary contexts. This approach to philosophical issues in the post-war period has recently been referred to by Marjorie Grene (1995: 55) as the "Bertie Wooster season in philosophy," a characterization I suspect would seem apt to many philosophers of mind interested in contemporary cognitive science (and in P. G. Wodehouse). Let me illustrate how this approach to philosophy served to isolate the philosophy of mind from the sciences of the mind with perhaps the two most influential examples pertaining to the mind in the ordinary language tradition.

In *The Concept of Mind,* Gilbert Ryle (1949: 17) attacked a view of the mind that he referred to as "Descartes' Myth" and "the dogma of the Ghost in the Machine"—basically, dualism—largely through a repeated application of the objection that dualism consisted of an extended *category mistake*: it "represents the facts of mental life as if they belonged to one logical type or category . . . when they actually belong to another." Descartes' Myth represented a category mistake because in supposing that there was a special, inner theater on which mental life is played out, it treated the "facts of mental life" as belonging to a special category of facts, when they were simply facts about how people can, do, and would behave in certain circumstances. Ryle set about showing that for the range of mental concepts that were held to refer to private, internal mental episodes or events according to Descartes' Myth—intelligence, the will, emotion, self-knowledge, sensation, and imagination—an appeal to what one would ordinarily say both shows the dogma of the Ghost in the Machine to be false, and points to a positive account of the mind that was behaviorist in orientation. To convey why Ryle's influential views here turned philosophy of mind away from science rather than towards it, consider the opening sentences of *The Concept of Mind*: "This book offers what may with reservations be described as a theory of the mind. But it does not give new information about minds. We possess already a wealth of information about minds, information which is neither derived from, nor upset by, the arguments of philosophers. The philosophical arguments which constitute this book are intended not to increase what we know about minds, but to rectify the logical geography of the knowledge which we already possess" (Ryle 1949: 9). The "we" here refers to ordinary folk, and the philosopher's task in articulating a theory of mind is to draw on what we already know about the mind, rather than on arcane, philosophical views or on specialized, scientific knowledge.

The second example is Norman Malcolm's *Dreaming,* which, like *The Concept of Mind,* framed the critique it wished to deliver as an attack on a Cartesian view of the mind. Malcolm's (1959: 4) target was the view that "dreams are the activity of the mind during sleep," and associated talk of DREAMING as involving various mental acts, such as remembering, imagining, judging, thinking, and reasoning. Malcolm argued that such dream-talk, whether it be part of commonsense reflection on dreaming (How long do dreams last?; Can you work out problems in your dreams?) or a contribution to more systematic empirical research on dreaming, was a confusion arising from the failure to attend to the proper "logic" of our ordinary talk about dreaming. Malcolm's argument proceeded by appealing to how one would *use* various expressions and sentences that contained the word "dreaming." (In looking back at Malcolm's book, it is striking that nearly every one of the eighteen short chapters begins with a paragraph about words and what one would say with or about them.)

Malcolm's central point was that there was no way to *verify* any given claim about such mental activity occurring while one was asleep, because the commonsense criteria for the application of such concepts were incompatible with saying that a person was asleep or dreaming. And because there was no way to tell whether various attributions of mental states to a sleeping person were correct, such attributions were meaningless. These claims not only could be made without an appeal to any empirical details about dreaming or SLEEP, but implied that the whole enterprise of investigating dreaming empirically itself represented some sort of *logical* muddle.

Malcolm's point became more general than one simply about dreaming (or the word "dreaming"). As he said in a preface to a later work, written after "the notion that thoughts, ideas, memories, sensations, and so on 'code into' or 'map onto' neural firing patterns in the brain" had become commonplace: "I believe that a study of our psychological concepts can show that [such] psycho-physical isomorphism is not a coherent assumption" (Malcolm 1971: x). Like Ryle's straightening of the logical geography of our knowledge of minds, Malcolm's appeal to the study of our psychological concepts could be conducted without any knowledge gleaned from psychological science (cf. Griffiths 1997: chap. 2 on the emotions).

Quite distinct from the ordinary language tradition was a second general perspective that served to make philosophical contributions to the study of the mind "distinctive" from those of science. This was logical positivism or empiricism, which developed in Europe in the 1920s and flourished in the United States through the 1930s and 1940s with the immigration to the United States of many of its leading members, including Rudolph Carnap, Hans Reichenbach, Herbert Feigl, and Carl Hempel. The logical empiricists were called "empiricists" because they held that it was via the senses and observation that we came to know about the world, deploying this empiricism with the logical techniques that had been developed by Gottlob Frege, Bertrand Russell, and Alfred Whitehead. Like empiricists in general, the logical positivists viewed the sciences as the paradigmatic repository of knowledge, and they were largely responsible for the rise of philosophy of science as a distinct subdiscipline within philosophy.

As part of their reflection on science they articulated and defended the doctrine of the UNITY OF SCIENCE, the idea that the sciences are, in some sense, essentially unified, and their empiricism led them to appeal to PARSIMONY AND SIMPLICITY as grounds for both theory choice within science and for preferring theories that were ontological Scrooges. This empiricism came with a focus on *what could be verified,* and with it scepticism about traditional metaphysical notions, such as God, CAUSATION, and essences, whose instances could not be verified by an appeal to the data of sense experience. This emphasis on verification was encapsulated in the verification theory of meaning, which held that the meaning of a sentence was its method of verification, implying that sentences without any such method were *meaningless.* In psychology, this fueled skepticism about the existence of internal mental representations and states (whose existence could not be objectively verified), and offered further philosophical backing for behaviorism.

In contrast to the ordinary language philosophers (many of whom would have been professionally embarrassed to have been caught knowing anything about science), the positivists held that philosophy was to be informed about and sensitive to the results of science. The distinctive task of the philosopher, however, was not simply to describe scientific practice, but to offer a *rational reconstruction* of it, one that made clear the logical structure of science. Although the term "*rational reconstruction*" was used first by Carnap in his 1928 book *The Logical Construction of the World,* quite a general epistemological tract, the technique to which it referred came to be applied especially to scientific concepts and theories.

This played out in the frequent appeal to the distinction between the *context of discovery* and the *context of justification,* drawn as such by Reichenbach in *Experience and Prediction* (1938) but with a longer history in the German tradition. To consider an aspect of a scientific view in the context of discovery was essentially to raise psychological, sociological, or historical questions about how that view originated, was developed, or came to be accepted or rejected. But properly philosophical explorations of science were to be conducted in the context of justification, raising questions and making claims about the logical structure of science and the concepts it used. Rational reconstruction was the chief way of divorcing the relevant scientific theory from its mere context of discovery.

A story involving Feigl and Carnap nicely illustrates the divorce between philosophy and science within positivism. In the late 1950s, Feigl visited the University of California, Los Angeles, to give a talk to the Department of Philosophy, of which Carnap was a member. Feigl's talk was aimed at showing that a form of physicalism, the

mind-brain identity theory, faced an empirical problem, since science had little, if anything, to say about the "raw feel" of consciousness, the WHAT-IT'S-LIKE of experience. During the question period, Carnap raised his hand, and was called on by Feigl. "Your claim that current neurophysiology tells us nothing about raw feels is wrong! You have overlooked the discovery of alpha-waves in the brain," exclaimed Carnap. Feigl, who was familiar with what he thought was the relevant science, looked puzzled: "Alpha-waves? What are they?" Carnap replied: "My dear Herbert. You tell me what raw feels are, and I will tell you what alpha-waves are."

Of the multiple readings that this story invites (whose common denominator is surely Carnap's savviness and wit), consider those that take Carnap's riposte to imply that he thought that one could defend materialism by, effectively, making up the science to fit whatever phenomena critics could rustle up. A rather extreme form of rational reconstruction, but it suggests one way in which the positivist approach to psychology could be just as a priori and so divorced from empirical practice as that of Ryle and Malcolm.

See also CAUSATION; DREAMING; PARSIMONY AND SIMPLICITY; SLEEP; UNITY OF SCIENCE; WHAT-IT'S-LIKE

4 *The Philosophy of Science*

The philosophy of science is integral to the cognitive sciences in a number of ways. We have already seen that positivists held views about the overall structure of science and the grounds for theory choice in science that had implications for psychology. Here I focus on three functions that the philosophy of science plays vis-à-vis the cognitive sciences: it provides a perspective on the place of psychology among the sciences; it raises questions about what any science can tell us about the world; and it explores the nature of knowledge and how it is known. I take these in turn.

One classic way in which the sciences were viewed as being unified, according to the positivists, was via reduction. REDUCTIONISM, in this context, is the view that intuitively "higher-level" sciences can be reduced, in some sense, to "lower-level" sciences. Thus, to begin with the case perhaps of most interest to MITECS readers, psychology was held to be reducible in principle to biology, biology to chemistry, chemistry to physics. This sort of reduction presupposed the existence of *bridge laws,* laws that exhaustively characterized the concepts of any higher-level science, and the generalizations stated using them, in terms of those concepts and generalizations at the next level down. And because reduction was construed as relating theories of one science to those of another, the advocacy of reductionism went hand-in-hand with a view of EXPLANATION that gave lower-level sciences at least a usurpatory power over their higher-level derivatives.

This view of the structure of science was opposed to EMERGENTISM, the view that the properties studied by higher-level sciences, such as psychology, were not mere aggregates of properties studied by lower-level sciences, and thus could not be completely understood in terms of them. Both emergentism and this form of reductionism were typically cast in terms of the relationship between laws in higher- and lower-level sciences, thus presupposing that there were, in the psychological case, PSYCHOLOGICAL LAWS in the first place. One well-known position that denies this assumption is Donald Davidson's ANOMALOUS MONISM, which claims that while mental states *are* strictly identical with physical states, our descriptions of them as mental states are neither definitionally nor nomologically reducible to descriptions of them as physical states. This view is usually expressed as denying the possibility of the bridge laws required for the reduction of psychology to biology.

Corresponding to the emphasis on scientific laws in views of the relations between the sciences is the idea that these laws state relations between NATURAL KINDS. The idea of a natural kind is that of a type or kind of thing that exists in the world itself, rather than a kind or grouping that exists because of our ways of perceiving, thinking about, or interacting with the world. Paradigms of natural kinds are biological kinds—species, such as the domestic cat (*Felis domesticus*)—and

chemical kinds—such as silver (Ag) and gold (Au). Natural kinds can be contrasted with *artifactual* kinds (such as chairs), whose members are artifacts that share common functions or purposes relative to human needs or designs; with *conventional* kinds (such as marriage vows), whose members share some sort of conventionally determined property; and from purely arbitrary groupings of objects, whose members have nothing significant in common save that they belong to the category. Views of what natural kinds are, of how extensively science traffics in them, and of how we should characterize the notion of a natural kind vis-à-vis other metaphysic notions, such as essence, intrinsic property, and causal power, all remain topics of debate in contemporary philosophy of science (e.g., van Fraassen 1989; Wilson 1999).

There is an intuitive connection between the claims that there are natural kinds, and that the sciences strive to identify them, and *scientific realism,* the view that the entities in mature sciences, whether they are observable or not, exist and our theories about them are at least approximately true. For realists hold that the sciences strive to "carve nature at its joints," and natural kinds are the pre-existing joints that one's scientific carving tries to find. The REALISM AND ANTIREALISM issue is, of course, more complicated than suggested by the view that scientific realists think there are natural kinds, and antirealists deny this—not least because there are a number of ways to deny either this realist claim or to diminish its significance. But such a perspective provides one starting point for thinking about the different views one might have of the relationship between science and reality.

Apart from raising issues concerning the relationships between psychology and other sciences and their respective objects of study, and questions about the relation between science and reality, the philosophy of science is also relevant to the cognitive sciences as a branch of epistemology or the theory of knowledge, studying a particular type of knowledge, scientific knowledge. A central notion in the general theory of knowledge is JUSTIFICATION, because being justified in what we believe is at least one thing that distinguishes knowledge from mere belief or a lucky guess. Since scientific knowledge is a paradigm of knowledge, views of justification have often been developed with scientific knowledge in mind.

The question of what it is for an individual to have a justified belief, however, has remained contentious in the theory of knowledge. Justified beliefs are those that we are entitled to hold, ones for which we have reasons, but how should we understand such entitlement and such reasons? One dichotomy here is between *internalists* about justification, who hold that having justified belief exclusively concerns facts that are "internal" to the believer, facts about his or her internal cognitive economy; and *externalists* about justification, who deny this. A second dichotomy is between *naturalists,* who hold that what cognitive states are justified may depend on facts about cognizers or about the world beyond cognizers that are uncovered by empirical science; and *rationalists,* who hold that justification is determined by the relations between one's cognitive states that the agent herself is in a special position to know about. Clearly part of what is at issue between internalists and externalists, as well as between naturalists and rationalists, is the role of the first-person perspective in accounts of justification and thus knowledge (see also Goldman 1997).

These positions about justification raise some general questions about the relationship between EPISTEMOLOGY AND COGNITION, and interact with views of the importance of first- and third-person perspectives on cognition itself. They also suggest different views of RATIONAL AGENCY, of what it is to be an agent who acts on the basis of justified beliefs. Many traditional views of rationality imply that cognizers have LOGICAL OMNISCIENCE, that is, that they believe all the logical consequences of their beliefs. Since clearly we are not logically omniscient, there is a question of how to modify one's account of rationality to avoid this result.

See also ANOMALOUS MONISM; EMERGENTISM; EPISTEMOLOGY AND COGNITION; EXPLANATION; JUSTIFICATION; LOGICAL OMNISCIENCE, PROBLEM OF; NATURAL KINDS; PSYCHOLOGICAL LAWS; RATIONAL AGENCY; REALISM AND ANTIREALISM; REDUCTIONISM

5 *The Mind in Cognitive Science*

At the outset, I said that the relation between the mental and physical remains the central, general issue in contemporary, materialist philosophy of mind. In section 2, we saw that the behaviorist critiques of Cartesian views of the mind and behaviorism themselves introduced a dilemma that derived from the problem of the homunculus that any mental science would seem to face. And in section 3 I suggested how a vibrant skepticism about the scientific status of a distinctively psychological science and philosophy's contribution to it was sustained by two dominant philosophical perspectives. It is time to bring these three points together as we move to explore the view of the mind that constituted the core of the developing field of cognitive science in the 1970s, what is sometimes called *classic* cognitive science, as well as its successors.

If we were to pose questions central to each of these three issues—the mental-physical relation, the problem of the homunculus, and the possibility of a genuinely cognitive science, they might be:

a. What is the relation between the mental and the physical?
b. How can psychology avoid the problem of the homunculus?
c. What makes a genuinely *mental* science possible?

Strikingly, these questions received standard answers, in the form of three "isms," from the nascent naturalistic perspective in the philosophy of mind that accompanied the rise of classic cognitive science. (The answers, so you don't have to peek ahead, are, respectively, functionalism, computationalism, and representationalism.)

The answer to (a) is FUNCTIONALISM, the view, baldly put, that mental states are functional states. Functionalists hold that what really matters to the identity of types of mental states is not what their instances are made of, but how those instances are causally arranged: what causes them, and what they, in turn, cause. Functionalism represents a view of the mental-physical relation that is compatible with materialism or physicalism because even if it is the functional or causal *role* that makes a mental state the state it is, every *occupant* of any particular role could be physical. The role-occupant distinction, introduced explicitly by Armstrong (1968) and implicitly in Lewis (1966), has been central to most formulations of functionalism.

A classic example of something that is functionally identified or individuated is *money:* it's not what it's made of (paper, gold, plastic) that makes something money but, rather, the causal role that it plays in some broader economic system. Recognizing this fact about money is not to give up on the idea that money is material or physical. Even though material composition is not what determines whether something is money, every instance of money is material or physical: dollar bills and checks are made of paper and ink, coins are made of metal, even money that is stored solely as a string of digits in your bank account has *some* physical composition. There are at least two related reasons why functionalism *about the mind* has been an attractive view to philosophers working in the cognitive sciences.

The first is that functionalism at least appears to support the AUTONOMY OF PSYCHOLOGY, for it claims that even if, as a matter of fact, our psychological states are realized in states of our brains, their status as *psychological* states lies in their functional organization, which can be abstracted from this particular material stuff. This is a *nonreductive* view of psychology. If functionalism is true, then there will be distinctively psychological natural kinds that cross-cut the kinds that are determined by a creature's material composition. In the context of materialism, functionalism suggests that creatures with very different material organizations could not only have mental states, but have *the same kinds* of mental states. Thus functionalism makes sense of comparative psychological or neurological investigations across species.

The second is that functionalism allows for *nonbiological* forms of intelligence and mentality. That is, because it is the "form" not the "matter" that determines psychological kinds, there could be entirely artifactual creatures, such as robots or computers, with mental states, provided that they have the right functional organization. This idea has been central to traditional artificial intelligence (AI), where one ideal has

been to create programs with a functional organization that not only allows them to behave in some crude way like intelligent agents but to do so in a way that instantiates at least some aspects of intelligence itself.

Both of these ideas have been criticized as part of attacks on functionalism. For example, Paul and Patricia Churchland (1981) have argued that the "autonomy" of psychology that one gains from functionalism can be a cover for the emptiness of the science itself, and Jaegwon Kim (1993) has argued against the coherence of the nonreductive forms of materialism usually taken to be implied by functionalism. Additionally, functionalism and AI are the targets of John Searle's much-discussed CHINESE ROOM ARGUMENT.

Consider (c), the question of what makes a distinctively mental science possible. Although functionalism gives one sort of answer to this in its basis for a defense of the autonomy (and so distinctness) of psychology, because there are more functional kinds than those in psychology (assuming functionalism), this answer does not explain what is distinctively *psychological* about psychology. A better answer to this question is *representationalism,* also known as the representational theory of mind. This is the view that mental states are relations between the bearers of those states and internal mental representations. Representationalism answers (c) by viewing psychology as the science concerned with the forms these mental representations can take, the ways in which they can be manipulated, and how they interact with one another in mediating between perceptual input and behavioral output.

A traditional version of representationalism, one cast in terms of Ideas, themselves often conceptualized as images, was held by the British empiricists John Locke, George Berkeley, and DAVID HUME. A form of representationalism, the LANGUAGE OF THOUGHT (LOT) hypothesis, has more recently been articulated and defended by Jerry Fodor (1975, 1981, 1987, 1994). The LOT hypothesis is the claim that we are able to cognize in virtue of having a mental language, *mentalese,* whose symbols are combined systematically by syntactic rules to form more complex units, such as thoughts. Because these mental symbols are intentional or representational (they are about things), the states that they compose are representational; mental states inherit their intentionality from their constituent mental representations.

Fodor himself has been particularly exercised to use the language of thought hypothesis to chalk out a place for the PROPOSITIONAL ATTITUDES and our folk psychology within the developing sciences of the mind. Not all proponents of the representational theory of mind, however, agree with Fodor's view that the system of representation underlying thought is a *language,* nor with his defense of folk psychology. But even forms of representationalism that are less committal than Fodor's own provide an answer to the question of what is distinctive about psychology: psychology is not mere neuroscience because it traffics in a range of mental representations and posits internal processes that operate on these representations.

Representationalism, particularly in Fodoresque versions that see the language of thought hypothesis as forming the foundations for a defense of both cognitive psychology and our commonsense folk psychology, has been challenged within cognitive science by the rise of connectionism in psychology and NEURAL NETWORKS within computer science. Connectionist models of psychological processing might be taken as an existence proof that one does not need to assume what is sometimes called the RULES AND REPRESENTATIONS approach to understand cognitive functions: the language of thought hypothesis is no longer "the only game in town."

Connectionist COGNITIVE MODELING of psychological processing, such as that of the formation of past tense (Rumelhart and McClelland 1986), face recognition (Cottrell and Metcalfe 1991), and VISUAL WORD RECOGNITION (Seidenberg and McClelland 1989), typically does not posit discrete, decomposable representations that are concatenated through the rules of some language of thought. Rather, connectionists posit a COGNITIVE ARCHITECTURE made up of simple neuron-like nodes, with activity being propagated across the units proportional to the weights of the connection strength between them. Knowledge lies not in the nodes themselves but in the values of the weights connecting nodes. There seems to be nothing of a propositional form within such connectionist networks, no place for the internal sentences that are the

objects of folk psychological states and other subpersonal psychological states posited in accounts of (for example) memory and reasoning.

The tempting idea that "classicists" accept, and connectionists reject, representationalism is too simple, one whose implausibility is revealed once one shifts one's focus from folk psychology and the propositional attitudes to cognition more generally. Even when research in classical cognitive science—for example, that on KNOWLEDGE-BASED SYSTEMS and on BAYESIAN NETWORKS—is cast in terms of "beliefs" that a system has, the connection between "beliefs" and the beliefs of folk psychology has been underexplored. More importantly, the notion of representation itself has not been abandoned across-the-board by connectionists, some of whom have sought to salvage and adapt the notion of mental representation, as suggested by the continuing debate over DISTRIBUTED VS. LOCAL REPRESENTATION and the exploration of sub-symbolic forms of representation within connectionism (see Boden 1990; Haugeland 1997; Smolensky 1994).

What perhaps better distinguishes classic and connectionist cognitive science here is not the issue of whether some form of representationalism is true, but whether the question to which it is an answer needs answering at all. In classical cognitive science, what makes the idea of a genuinely *mental* science possible is the idea that psychology describes representation crunching. But in starting with the idea that neural representation occurs from single neurons up through circuits to modules and more nebulous, distributed neural systems, connectionists are less likely to think that psychology offers a distinctive level of explanation that deserves some identifying characterization. This rejection of question (c) is clearest, I think, in related DYNAMIC APPROACHES TO COGNITION, since such approaches investigate psychological states as dynamic systems that need not posit distinctly *mental* representations. (As with connectionist theorizing about cognition, dynamic approaches encompass a variety of views of mental representation and its place in the study of the mind that make representationalism itself a live issue within such approaches; see Haugeland 1991; van Gelder 1998.)

Finally, consider (b), the question of how to avoid the problem of the homunculus in the sciences of the mind. In classic cognitive science, the answer to (b) is *computationalism,* the view that mental states are computational, an answer which integrates and strengthens functionalist materialism and representationalism as answers to our previous two questions. It does so in the *way* in which it provides a more precise characterization of the nature of the functional or causal relations that exist between mental states: these are *computational relations between mental representations.* The traditional way to spell this out is the COMPUTATIONAL THEORY OF MIND, according to which the mind is a digital computer, a device that stores symbolic representations and performs operations on them in accord with *syntactic* rules, rules that attend only to the "form" of these symbols. This view of computationalism has been challenged not only by relatively technical objections (such as that based on the FRAME PROBLEM), but also by the development of neural networks and models of SITUATED COGNITION AND LEARNING, where (at least some) informational load is shifted from internal codes to organism-environment interactions (cf. Ballard et al. 1997).

The computational theory of mind avoids the problem of the homunculus because digital computers that exhibit some intelligence exist, and they do not contain undischarged homunculi. Thus, if *we* are fancy versions of such computers, then we can understand our intelligent capacities without positing undischarged homunculi. The way this works in computers is by having a series of programs and languages, each compiled by the one beneath it, with the most basic language directly implemented in the hardware of the machine. We avoid an endless series of homunculi because the capacities that are posited at any given level are typically simpler and more numerous than those posited at any higher level, with the lowest levels specifying instructions to perform actions that require no intelligence at all. This strategy of FUNCTIONAL DECOMPOSITION solves the problem of the homunculus if we are digital computers, assuming that it solves it for digital computers.

Like representationalism, computationalism has sometimes been thought to have been superseded by either (or both) the connectionist revolution of the 1980s, or the

Decade of the Brain (the 1990s). But as with proclamations of the death of representationalism, this notice of the death of computationalism is premature. In part this is because the object of criticism is a specific version of computationalism, not computationalism per se (cf. representationalism), and in part it is because neural networks and the neural systems in the head they model are both themselves typically claimed to be computational in some sense. It is surprisingly difficult to find an answer within the cognitive science community to the question of whether there is a univocal notion of COMPUTATION that underlies the various different computational approaches to cognition on offer. The various types of AUTOMATA postulated in the 1930s and 1940s—particularly TURING machines and the "neurons" of MCCULLOCH and PITTS, which form the intellectual foundations, respectively, for the computational theory of mind and contemporary neural network theory—have an interwoven history, and many of the initial putative differences between classical and connectionist cognitive science have faded into the background as research in artificial intelligence and cognitive modeling has increasingly melded the insights of each approach into more sophisticated hybrid models of cognition (cf. Ballard 1997).

While dynamicists (e.g., Port and van Gelder 1995) have sometimes been touted as providing a noncomputational alternative to both classic and connectionist cognitive science (e.g., Thelen 1995: 70), as with claims about the nonrepresentational stance of such approaches, such a characterization is not well founded (see Clark 1997, 1998). More generally, the relationship between dynamical approaches to both classical and connectionist views remains a topic for further discussion (cf. van Gelder and Port 1995; Horgan and Tienson 1996; and Giunti 1997).

See also AUTOMATA; AUTONOMY OF PSYCHOLOGY; BAYESIAN NETWORKS; CHINESE ROOM ARGUMENT; COGNITIVE ARCHITECTURE; COGNITIVE MODELING, CONNECTIONIST; COGNITIVE MODELING, SYMBOLIC; COMPUTATION; COMPUTATIONAL THEORY OF MIND; DISTRIBUTED VS. LOCAL REPRESENTATION; DYNAMIC APPROACHES TO COGNITION; FRAME PROBLEM; FUNCTIONAL DECOMPOSITION; FUNCTIONALISM; HUME, DAVID; KNOWLEDGE-BASED SYSTEMS; LANGUAGE OF THOUGHT; MCCULLOCH, WARREN S.; NEURAL NETWORKS; PITTS, WALTER; PROPOSITIONAL ATTITUDES; RULES AND REPRESENTATIONS; SITUATED COGNITION AND LEARNING; TURING, ALAN; VISUAL WORD RECOGNITION

6 *A Focus on Folk Psychology*

Much recent philosophical thinking about the mind and cognitive science remains preoccupied with the three traditional philosophical issues I identified in the first section: the mental-physical relation, the structure of the mind, and the first-person perspective. All three issues arise in one of the most absorbing discussions over the last twenty years, that over the nature, status, and future of what has been variously called commonsense psychology, the propositional attitudes, or FOLK PSYCHOLOGY.

The term *folk psychology* was coined by Daniel Dennett (1981) to refer to the systematic knowledge that we "folk" employ in explaining one another's thoughts, feelings, and behavior; the idea goes back to Sellars's Myth of Jones in "Empiricism and the Philosophy of Mind" (1956). We all naturally and without explicit instruction engage in psychological explanation by attributing beliefs, desires, hopes, thoughts, memories, and emotions to one another. These patterns of folk psychological explanation are "folk" as opposed to "scientific" since they require no special training and are manifest in everyday predictive and explanatory practice; and genuinely "psychological" because they posit the existence of various states or properties that seem to be paradigmatically mental in nature. To engage in folk psychological explanation is, in Dennett's (1987) terms, to adopt the INTENTIONAL STANCE.

Perhaps the central issue about folk psychology concerns its relationship to the developing cognitive sciences. ELIMINATIVE MATERIALISM, or eliminativism, is the view that folk psychology will find no place in any of the sciences that could be called "cognitive" in orientation; rather, the fortune of folk psychology will be like that of many other folk views of the world that have found themselves permanently out of

step with scientific approaches to the phenomena they purport to explain, such as folk views of medicine, disease, and witchcraft.

Eliminativism is sometimes motivated by adherence to reductionism (including the thesis of EXTENSIONALITY) and the ideal of the unity of science, together with the recognition that the propositional attitudes have features that set them off in kind from the types of entity that exist in other sciences. For example, they are intentional or representational, and attributing them to individuals seems to depend on factors beyond the boundary of those individuals, as the TWIN EARTH arguments suggest. These arguments and others point to a prima facie conflict between folk psychology and INDIVIDUALISM (or *internalism*) in psychology (see Wilson 1995). The apparent conflict between folk psychology and individualism has provided one of the motivations for developing accounts of NARROW CONTENT, content that depends solely on an individual's intrinsic, physical properties. (The dependence here has usually been understood in terms of the technical notion of SUPERVENIENCE; see Horgan 1993.)

There is a spin on this general motivation for eliminative materialism that appeals more directly to the issue of the how the mind is structured. The claim here is that whether folk psychology is defensible will turn in large part on how compatible its ontology—its list of what we find in a folk psychological mind—is with the developing ontology of the cognitive sciences. With respect to classical cognitive science, with its endorsement of both the representational and computational theories of mind, folk psychology is on relatively solid ground here. It posits representational states, such as belief and desire, and it is relatively easy to see how the causal relations between such states could be modeled computationally. But connectionist models of the mind, with what representation there is lying in patterns of activity rather than in explicit representations like propositions, seem to leave less room in the structure of the mind for folk psychology.

Finally, the issue of the place of the first-person perspective arises with respect to folk psychology when we ask how people deploy folk psychology. That is, what sort of psychological machinery do we folk employ in engaging in folk psychological explanation? This issue has been the topic of the SIMULATION VS. THEORY-THEORY debate, with proponents of the simulation view holding, roughly, a "first-person first" account of how folk psychology works, and theory-theory proponents viewing folk psychology as essentially a third-person predictive and explanatory tool. Two recent volumes by Davies and Stone (1995a, 1995b) have added to the literature on this debate, which has developmental and moral aspects, including implications for MORAL PSYCHOLOGY.

See also ELIMINATIVE MATERIALISM; EXTENSIONALITY, THESIS OF; FOLK PSYCHOLOGY; INDIVIDUALISM; INTENTIONAL STANCE; MORAL PSYCHOLOGY; NARROW CONTENT; SIMULATION VS. THEORY-THEORY; SUPERVENIENCE; TWIN EARTH

7 *Exploring Mental Content*

Although BRENTANO's claim that INTENTIONALITY is the "mark of the mental" is problematic and has few adherents today, intentionality has been one of the flagship topics in philosophical discussion of the mental, and so at least a sort of mark of that discussion. Just what the puzzle about intentionality is and what one might say about it are topics I want to explore in more detail here.

To say that something is intentional is just to say that it is *about something*, or that it *refers to something*. In this sense, statements of fact are paradigmatically intentional, since they are about how things are in the world. Similarly, a highway sign with a picture of a gas pump on it is intentional because it conveys the information that there is gas station ahead at an exit: it is, in some sense, about that state of affairs.

The beginning of chapter 4 of Jerry Fodor's *Psychosemantics* provides one lively expression of the problem with intentionality:

> I suppose that sooner or later the physicists will complete the catalogue they've been compiling of the ultimate and irreducible properties of things. When they do, the likes of *spin*, *charm*, and *charge* will perhaps appear upon their list. But *aboutness* surely won't; intentionality simply doesn't go that deep. It's hard to see, in face of this consideration, how one can be a Realist about intentionality without

> also being, to some extent or other, a Reductionist. If the semantic and the intentional are real properties of things, it must be in virtue of their identity with (or maybe of their supervenience on?) properties that are themselves *neither* intentional *nor* semantic. If aboutness is real, it must be really something else. (p. 97, emphases in original)

Although there is much that one could take issue with in this passage, my reason for introducing it here is not to critique it but to try to capture some of the worries about intentionality that bubble up from it.

The most general of these concerns the *basis* of intentionality in the natural order: given that only special parts of the world (like our minds) have intentional properties, what is it about those things that gives them (and not other things) intentionality? Since not only mental phenomena are intentional (for example, spoken and written natural language and systems of signs and codes are as well), one might think that a natural way to approach this question would be as follows. Consider all of the various sorts of "merely material" things that at least seem to have intentional properties. Then proceed to articulate why each of them is intentional, either taking the high road of specifying something like the "essence of intentionality"—something that all and only things with intentional properties have—or taking the low road of doing so for each phenomenon, allowing these accounts to vary across disparate intentional phenomena.

Very few philosophers have explored the problem of intentionality in this way. I think this is chiefly because they do not view all things with intentional properties as having been created equally. A common assumption is that even if lots of the nonmental world is intentional, its intentionality is *derived,* in some sense, from the intentionality of the mental. So, to take a classic example, the sentences we utter and write are intentional all right (they are about things). But their intentionality derives from that of the corresponding thoughts that are their causal antecedents. To take another oft-touted example, computers often produce intentional output (even photocopiers can do this), but whatever intentionality lies in such output is not inherent to the machines that produce it but is derivative, ultimately, from the mental states of those who design, program, and use them and their products. Thus, there has been a focus on mental states as a sort of paradigm of intentional state, and a subsequent narrowing of the sorts of intentional phenomena discussed. Two points are perhaps worth making briefly in this regard.

First, the assumption that not all things with intentional properties are created equally is typically shared even by those who have not focused almost exclusively on mental states as paradigms of intentional states, but on languages and other public and conventional forms of representation (e.g., Horst 1996). It is just that their paradigm is different.

Second, even when mental states *have* been taken as a paradigm here, those interested in developing a "psychosemantics"—an account of the basis for the semantics of psychological states—have often turned to decidedly nonmental systems of representation in order to theorize about the intentionality of the mental. This focus on what we might think of as *proto-intentionality* has been prominent within both Fred Dretske's (1981) informational semantics and the biosemantic approach pioneered by Ruth Millikan (1984, 1993).

The idea common to such views is to get clear about the grounds of simple forms of intentionality before scaling up to the case of the intentionality of human minds, an instance of a research strategy that has driven work in the cognitive sciences from early work in artificial intelligence on KNOWLEDGE REPRESENTATION and cognitive modeling through to contemporary work in COMPUTATIONAL NEUROSCIENCE. Exploring simplified or more basic intentional systems in the hope of gaining some insight into the more full-blown case of the intentionality of human minds runs the risk, of course, of focusing on cases that leave out precisely that which is crucial to full-blown intentionality. Some (for example, Searle 1992) would claim that consciousness and phenomenology are such features.

As I hinted at in my discussion of the mind in cognitive science in section 5, construed one way the puzzle about the grounds of intentionality has a general answer in the hypothesis of computationalism. But there is a deeper problem about the grounds

of intentionality concerning *just how* at least some mental stuff could be about other stuff in the world, and computationalism is of little help here. Computationalism does not even pretend to answer the question of what it is about specific mental states (say, my belief that trees often have leaves) that gives them the content that they have—for example, that makes them *about trees*. Even if we *were* complicated Turing machines, what would it be about *my* Turing machine table that implies that I have the belief that trees often have leaves? Talking about the correspondence between the semantic and syntactic properties that symbol structures in computational systems have, and of how the former are "inherited" from the latter is well and good. But it leaves open the "just how" question, and so fails to address what I am here calling the deeper problem about the grounds of intentionality. This problem is explored in the article on MENTAL REPRESENTATION, and particular proposals for a psychosemantics can be found in those on INFORMATIONAL SEMANTICS and FUNCTIONAL ROLE SEMANTICS.

It would be remiss in exploring mental content to fail to mention that much thought about intentionality has been propelled by work in the philosophy of language: on INDEXICALS AND DEMONSTRATIVES, on theories of REFERENCE and the propositional attitudes, and on the idea of RADICAL INTERPRETATION. Here I will restrict myself to some brief comments on theories of reference, which have occupied center stage in the philosophy of language for much of the last thirty years.

One of the central goals of theories of reference has been to explain in virtue of what parts of sentences of natural languages refer to the things they refer to. What makes the name "Miranda" refer to my daughter? In virtue of what does the plural noun "dogs" refer to dogs? Such questions have a striking similarity to my above expression of the central puzzle concerning intentionality. In fact, the application of causal theories of reference (Putnam 1975, Kripke 1980) developed principally for natural languages has played a central role in disputes in the philosophy of mind that concern intentionality, including those over individualism, narrow content, and the role of Twin Earth arguments in thinking about intentionality. In particular, applying them not to the meaning of natural language terms but to the content of thought is one way to reach the conclusion that *mental* content does not supervene on an individual's physical properties, that is, that mental content is not individualistic.

GOTTLOB FREGE is a classic source for contrasting descriptivist theories of reference, according to which natural language reference is, in some sense, mediated by a speaker's descriptions of the object or property to which she refers. Moreover, Frege's notion of sense and the distinction between SENSE AND REFERENCE are often invoked in support of the claim that there is much to MEANING—linguistic or mental—that goes beyond the merely referential. Frege is also one of the founders of modern logic, and it is to the role of logic in the cognitive sciences that I now turn.

See also BRENTANO, FRANZ; COMPUTATIONAL NEUROSCIENCE; FREGE, GOTTLOB; FUNCTIONAL ROLE SEMANTICS; INDEXICALS AND DEMONSTRATIVES; INFORMATIONAL SEMANTICS; INTENTIONALITY; KNOWLEDGE REPRESENTATION; MEANING; MENTAL REPRESENTATION; RADICAL INTERPRETATION; REFERENCE, THEORIES OF; SENSE AND REFERENCE

8 *Logic and the Sciences of the Mind*

Although INDUCTION, like deduction, involves drawing inferences on the basis of one or more premises, it is *deductive* inference that has been the focus in LOGIC, what is often simply referred to as "formal logic" in departments of philosophy and linguistics. The idea that it is possible to abstract away from deductive arguments given in natural language that differ in the content of their premises and conclusions goes back at least to Aristotle in the fourth century B.C. Hence the term "Aristotelian syllogisms" to refer to a range of argument forms containing premises and conclusions that begin with the words "every" or "all," "some," and "no." This abstraction makes it possible to talk about argument *forms* that are valid and invalid, and allows one to describe two arguments as being of the same *logical* form. To take a simple example, we know that any argument of the form:

All A are B.
No B are C.

No A are C.

is *formally* valid, where the emphasis here serves to highlight reference to the preservation of truth from premises to conclusion, that is, the validity, solely in virtue of the forms of the individual sentences, together with the form their arrangement constitutes. Whatever plural noun phrases we substitute for "A," "B," and "C," the resulting natural language argument will be valid: if the two premises are true, the conclusion must also be true. The same general point applies to arguments that are formally *invalid,* which makes it possible to talk about formal *fallacies,* that is, inferences that are invalid because of the forms they instantiate.

Given the age of the general idea of LOGICAL FORM, what is perhaps surprising is that it is only in the late nineteenth century that the notion was developed so as to apply to a wide range of natural language constructions through the development of the *propositional* and *predicate* logics. And it is only in the late twentieth century that the notion of logical form comes to be appropriated within linguistics in the study of SYNTAX. I focus here on the developments in logic.

Central to propositional logic (sometimes called "sentential logic") is the idea of a propositional or sentential *operator,* a symbol that acts as a function on propositions or sentences. The paradigmatic propositional operators are symbols for negation ("~"), conjunction ("&"), disjunction ("v"), and conditional ("→"). And with the development of formal languages containing these symbols comes an ability to represent a richer range of formally valid arguments, such as that manifest in the following thought:

> If Sally invites Tom, then either he will say "no," or cancel his game with Bill. But there's no way he'd turn Sally down. So I guess if she invites him, he'll cancel with Bill.

In predicate or quantificational logic, we are able to represent not simply the relations between propositions, as we can in propositional logic, but also the structure within propositions themselves through the introduction of QUANTIFIERS and the terms and predicates that they bind. One of the historically more important applications of predicate logic has been its widespread use in linguistics, philosophical logic, and the philosophy of language to formally represent increasingly larger parts of natural languages, including not just simple subjects and predicates, but adverbial constructions, tense, indexicals, and attributive adjectives (for example, see Sainsbury 1991).

These fundamental developments in logical theory have had perhaps the most widespread and pervasive effect on the foundations of the cognitive sciences of *any* contributions from philosophy or mathematics. They also form the basis for much contemporary work across the cognitive sciences: in linguistic semantics (e.g., through MODAL LOGIC, in the use of POSSIBLE WORLDS SEMANTICS to model fragments of natural language, and in work on BINDING); in metalogic (e.g., on FORMAL SYSTEMS and results such as the CHURCH-TURING THESIS and GÖDEL'S THEOREMS); and in artificial intelligence (e.g., on LOGICAL REASONING SYSTEMS, TEMPORAL REASONING, and METAREASONING).

Despite their technical payoff, the relevance of these developments in logical theory for thinking more directly about DEDUCTIVE REASONING in human beings is, ironically, less clear. Psychological work on human reasoning, including that on JUDGMENT HEURISTICS, CAUSAL REASONING, and MENTAL MODELS, points to ways in which human reasoning may be governed by structures very different from those developed in formal logic, though this remains an area of continuing debate and discussion.

See also BINDING THEORY; CAUSAL REASONING; CHURCH-TURING THESIS; DEDUCTIVE REASONING; FORMAL SYSTEMS, PROPERTIES OF; GÖDEL'S THEOREMS; INDUCTION; JUDGMENT HEURISTICS; LOGIC; LOGICAL FORM IN LINGUISTICS; LOGICAL FORM, ORIGINS OF; LOGICAL REASONING SYSTEMS; MENTAL MODELS; METAREASONING;

MODAL LOGIC; POSSIBLE WORLDS SEMANTICS; QUANTIFIERS; SYNTAX; TEMPORAL REASONING

9 *Two Ways to Get Biological*

By the late nineteenth century, both evolutionary theory and the physiological study of mental capacities were firmly entrenched. Despite this, these two paths to a biological view of cognition have only recently been re-explored in sufficient depth to warrant the claim that contemporary cognitive science incorporates a truly biological perspective on the mind. The neurobiological path, laid down by the tradition of physiological psychology that developed from the mid-nineteenth century, is certainly the better traveled of the two. The recent widening of this path by those dissatisfied with the distinctly nonbiological approaches adopted within traditional artificial intelligence has, as we saw in our discussion of computationalism, raised new questions about COMPUTATION AND THE BRAIN, the traditional computational theory of the mind, and the rules and representations approach to understanding the mind. The evolutionary path, by contrast, has been taken only occasionally and half-heartedly over the last 140 years. I want to concentrate not only on why but on the ways in which evolutionary theory is relevant to contemporary interdisciplinary work on the mind.

The theory of EVOLUTION makes a claim about the *patterns* that we find in the biological world—they are patterns of *descent*—and a claim about the predominant cause of those patterns—they are caused by the mechanism of natural selection. None of the recent debates concerning evolutionary theory—from challenges to the focus on ADAPTATION AND ADAPTATIONISM in Gould and Lewontin (1979) to more recent work on SELF-ORGANIZING SYSTEMS and ARTIFICIAL LIFE—challenges the substantial core of the theory of evolution (cf. Kauffman 1993, 1995; Depew and Weber 1995). The vast majority of those working in the cognitive sciences both accept the theory of evolution and so think that a large number of traits that organisms possess are adaptations to evolutionary forces, such as natural selection. Yet until the last ten years, the scattered pleas to apply evolutionary theory to the mind (such as those of Ghiselin 1969 and Richards 1987) have come largely from those outside of the psychological and behavioral sciences.

Within the last ten years, however, a distinctive EVOLUTIONARY PSYCHOLOGY has developed as a research program, beginning in Leda Cosmides's (1989) work on human reasoning and the Wason selection task, and represented in the collection of papers *The Adapted Mind* (Barkow, Cosmides, and Tooby 1992) and, more recently and at a more popular level, by Steven Pinker's *How the Mind Works* (1997). Evolutionary psychologists view the mind as a set of "Darwinian algorithms" designed by natural selection to solve adaptive problems faced by our hunter-gatherer ancestors. The claim is that this basic Darwinian insight can and should guide research into the cognitive architecture of the mind, since the task is one of discovering and understanding the *design* of the human mind, in all its complexity. Yet there has been more than an inertial resistance to viewing evolution as central to the scientific study of human cognition.

One reason is that evolutionary theory in general is seen as answering different questions than those at the core of the cognitive sciences. In terms of the well-known distinction between *proximal* and *ultimate* causes, appeals to evolutionary theory primarily allow one to specify the latter, and cognitive scientists are chiefly interested in the former: they are interested in the *how* rather than the *why* of the mind. Or to put it more precisely, central to cognitive science is an understanding of the *mechanisms* that govern cognition, not the various histories—evolutionary or not—that produced these mechanisms. This general perception of the concerns of evolutionary theory and the contrasting conception of cognitive science, have both been challenged by evolutionary psychologists. The same general challenges have been issued by those who think that the relations between ETHICS AND EVOLUTION and those between cognition and CULTURAL EVOLUTION have not received their due in contemporary cognitive science.

Yet despite the skepticism about this direct application of evolutionary theory to human cognition, its implicit application is inherent in the traditional interest in the

minds of *other* animals, from *aplysia* to (nonhuman) apes. ANIMAL NAVIGATION, PRIMATE LANGUAGE, and CONDITIONING AND THE BRAIN, while certainly topics of interest in their own right, gain some added value from what their investigation can tell us about *human* minds and brains. This presupposes something like the following: that there are natural kinds in psychology that transcend species boundaries, such that there is a general way of exploring how a cognitive capacity is structured, independent of the particular species of organism in which it is instantiated (cf. functionalism). Largely on the basis of research with non-human animals, we know enough now to say, with a high degree of certainty, things like this: that the CEREBELLUM is the central brain structure involved in MOTOR LEARNING, and that the LIMBIC SYSTEM plays the same role with respect to at least some EMOTIONS.

This is by way of returning to (and concluding with) the neuroscientific path to biologizing the mind, and the three classic philosophical issues about the mind with which we began. As I hope this introduction has suggested, despite the distinctively philosophical edge to all three issues—the mental-physical relation, the structure of the mind, and the first-person perspective—discussion of each of them is elucidated and enriched by the interdisciplinary perspectives provided by empirical work in the cognitive sciences. It is not only a priori arguments but complexities revealed by empirical work (e.g., on the neurobiology of consciousness, or ATTENTION and animal and human brains) that show the paucity of the traditional philosophical "isms" (dualism, behaviorism, type-type physicalism) with respect to the mental-physical relation. It is not simply general, philosophical arguments against nativism or against empiricism about the structure of the mind that reveal limitations to the global versions of these views, but ongoing work on MODULARITY AND LANGUAGE, on cognitive architecture, and on the innateness of language. And thought about introspection and self-knowledge, to take two topics that arise when one reflects on the first-person perspective on the mind, is both enriched by and contributes to empirical work on BLINDSIGHT, the theory of mind, and METAREPRESENTATION. With some luck, philosophers increasingly sensitive to empirical data about the mind will have paved a two-way street that encourages psychologists, linguists, neuroscientists, computer scientists, social scientists and evolutionary theorists to venture more frequently and more surely into philosophy.

See also ADAPTATION AND ADAPTATIONISM; ANIMAL NAVIGATION; ARTIFICIAL LIFE; ATTENTION IN THE ANIMAL BRAIN; ATTENTION IN THE HUMAN BRAIN; BLINDSIGHT; CEREBELLUM; COMPUTATION AND THE BRAIN; CONDITIONING AND THE BRAIN; CULTURAL EVOLUTION; EMOTIONS; ETHICS AND EVOLUTION; EVOLUTION; EVOLUTIONARY PSYCHOLOGY; LIMBIC SYSTEM; METAREPRESENTATION; MODULARITY AND LANGUAGE; MOTOR LEARNING; PRIMATE LANGUAGE; SELF-ORGANIZING SYSTEMS

Acknowledgments

I would like to thank Kay Bock, Bill Brewer, Alvin Goldman, John Heil, Greg Murphy, Stewart Saunders, Larry Shapiro, Sydney Shoemaker, Tim van Gelder, and Steve Wagner, as well as the PNP Group at Washington University, St. Louis, for taking time out to provide some feedback on earlier versions of this introduction. I guess the remaining idiosyncrasies and mistakes are mine.

References

Armstrong, D. M. (1968). *A Materialist Theory of the Mind.* London: Routledge and Kegan Paul.

Ballard, D. (1997). *An Introduction to Natural Computation.* Cambridge, MA: MIT Press.

Ballard, D., M. Hayhoe, P. Pook, and R. Rao. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences* 20: 723–767.

Barkow, J. H., L. Cosmides, and J. Tooby, Eds. (1992). *The Adapted Mind.* New York: Oxford University Press.

Boden, M., Ed. (1990). *The Philosophy of Artificial Intelligence.* Oxford: Oxford University Press.

Carnap, R. (1928). *The Logical Construction of the World.* Translated by R. George (1967). Berkeley: University of California Press.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory.* New York: Oxford University Press.

Chomsky, N. (1959). Review of B. F. Skinner's *Verbal Behavior. Language* 35 : 26–58.
Churchland, P. M. (1979). *Scientific Realism and the Plasticity of Mind.* New York: Cambridge University Press.
Churchland, P. M., and P. S. Churchland. (1981). Functionalism, qualia, and intentionality. *Philosophical Topics* 12: 121–145.
Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again.* Cambridge, MA: MIT Press.
Clark, A. (1998). Twisted tales: Causal complexity and cognitive scientific explanation. *Minds and Machines* 8: 79–99.
Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason Selection Task. *Cognition* 31: 187–276.
Cottrell, G., and J. Metcalfe. (1991). EMPATH: Face, Emotion, and Gender Recognition Using Holons. In R. Lippman, J. Moody, and D. Touretzky, Eds., *Advances in Neural Information Processing Systems,* vol. 3. San Mateo, CA: Morgan Kaufmann.
Davies, M., and T. Stone, Eds. (1995a). *Folk Psychology: The Theory of Mind Debate.* Oxford: Blackwell.
Davies, M., and T. Stone, Eds. (1995b). *Mental Simulation: Evaluations and Applications.* Oxford: Blackwell.
Dennett, D. C. (1981). Three kinds of intentional psychology. Reprinted in his 1987.
Dennett, D. C. (1987). *The Intentional Stance.* Cambridge, MA: MIT Press.
Depew, D., and B. Weber. (1995). *Darwinism Evolving: Systems Dynamics and the Genealogy of Natural Selection.* Cambridge, MA: MIT Press.
Dretske, F. (1981). *Knowledge and the Flow of Information.* Cambridge, MA: MIT Press.
Elman, J., E. Bates, M. Johnson, A. Karmiloff-Smith, D. Parisi, and K. Plunkett, Eds. (1996). *Rethinking Innateness.* Cambridge, MA: MIT Press.
Fodor, J. A. (1975). *The Language of Thought.* Cambridge, MA: Harvard University Press.
Fodor, J. A. (1981). *Representations: Philosophical Essays on the Foundations of Cognitive Science.* Sussex: Harvester Press.
Fodor, J. A. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* Cambridge, MA: MIT Press.
Fodor, J. A. (1994). *The Elm and the Expert.* Cambridge, MA: MIT Press.
Geach, P. (1957). *Mental Acts.* London: Routledge and Kegan Paul.
Ghiselin, M. (1969). *The Triumph of the Darwinian Method.* Berkeley: University of California Press.
Giunti, M. (1997). *Computation, Dynamics, and Cognition.* New York: Oxford University Press.
Goldman, A. (1997). Science, Publicity, and Consciousness. *Philosophy of Science* 64: 525–545.
Gould, S. J., and R. C. Lewontin. (1979). The spandrels of San Marco and the panglossian paradigm: A critique of the adaptationist programme. Reprinted in E. Sober, Ed., *Conceptual Issues in Evolutionary Biology,* 2nd ed. (1993.) Cambridge, MA: MIT Press.
Grene, M. (1995). *A Philosophical Testament.* Chicago: Open Court.
Griffiths, P. E. (1997). *What Emotions Really Are.* Chicago: University of Chicago Press.
Haugeland, J. (1991). Representational genera. In W. Ramsey and S. Stich, Eds., *Philosophy and Connectionist Theory.* Hillsdale, NJ: Erlbaum.
Haugeland, J., Ed. (1997). *Mind Design 2: Philosophy, Psychology, and Artificial Intelligence.* Cambridge, MA: MIT Press.
Heil, J., and A. Mele, Eds. (1993). *Mental Causation.* Oxford: Clarendon Press.
Horgan, T. (1993). From supervenience to superdupervenience: Meeting the demands of a material world. *Mind* 102: 555–586.
Horgan, T., and J. Tienson. (1996). *Connectionism and the Philosophy of Psychology.* Cambridge, MA: MIT Press.
Horst, S. (1996). *Symbols, Computation, and Intentionality.* Berkeley: University of California Press.
James, W. (1890). *The Principles of Psychology.* 2 vol. Dover reprint (1950). New York: Dover.
Kauffman, S. (1993). *The Origins of Order.* New York: Oxford University Press.
Kauffman, S. (1995). *At Home in the Universe.* New York: Oxford University Press.
Kim, J. (1993). *Supervenience and Mind.* New York: Cambridge University Press.
Kosslyn, S. (1980). *Image and Mind.* Cambridge, MA: Harvard University Press.
Kosslyn, S. (1994). *Image and Brain.* Cambridge, MA: MIT Press.
Kripke, S. (1980). *Naming and Necessity.* Cambridge, MA: Harvard University Press.
Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly* 64: 354–361.
Lewis, D. K. (1966). An argument for the identity theory. *Journal of Philosophy* 63: 17–25.
Malcolm, N. (1959). *Dreaming.* London: Routledge and Kegan Paul.
Malcolm, N. (1971). *Problems of Mind: Descartes to Wittgenstein.* New York: Harper and Row.
Millikan, R. G. (1984). *Language, Thought, and Other Biological Categories.* Cambridge, MA: MIT Press.
Millikan, R. G. (1993). *White Queen Psychology and Other Essays for Alice.* Cambridge, MA: MIT Press.

Pinker, S. (1997). *How the Mind Works.* New York: Norton.
Port, R., and T. van Gelder, Eds. (1995). *Mind as Motion: Explorations in the Dynamics of Cognition.* Cambridge, MA: MIT Press.
Putnam, H. (1975). The meaning of "meaning." Reprinted in *Mind, Language, and Reality: Collected Papers,* vol. 2. Cambridge: Cambridge University Press.
Pylyshyn, Z. (1984). *Computation and Cognition.* Cambridge, MA: MIT Press.
Reichenbach, H. (1938). *Experience and Prediction.* Chicago: University of Chicago Press.
Richards, R. (1987). *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior.* Chicago: University of Chicago Press.
Rumelhart, D., and J. McClelland. (1986). On Learning the Past Tenses of English Verbs. In J. McClelland, D. Rumelhart, and the PDP Research Group, Eds., *Parallel Distributed Processing,* vol. 2. Cambridge, MA: MIT Press.
Ryle, G. (1949). *The Concept of Mind.* New York: Penguin.
Sainsbury, M. (1991). *Logical Forms.* New York: Blackwell.
Searle, J. (1992). *The Rediscovery of the Mind.* Cambridge, MA: MIT Press.
Seidenberg, M. S., and J. L. McClelland. (1989). A distributed, developmental model of visual word recognition and naming. *Psychological Review* 96: 523–568.
Sellars, W. (1956). Empiricism and the philosophy of mind. In H. Feigl and M. Scriven, Eds., *Minnesota Studies in the Philosophy of Science,* vol. 1. Minneapolis: University of Minnesota Press.
Skinner, B. F. (1957). *Verbal Behavior.* New York: Appleton-Century-Crofts.
Smolensky, P. (1994). Computational models of mind. In S. Guttenplan, Ed., *A Companion to the Philosophy of Mind.* Cambridge, MA: Blackwell.
Spelke, E. (1990). Principles of object perception. *Cognitive Science* 14: 29–56.
Stabler, E. (1983). How are grammars represented? *Behavioral and Brain Sciences* 6: 391–420.
Thelen, E. (1995). Time-scale dynamics and the development of an embodied cognition. In R. Port and T. van Gelder, Eds., *Mind as Motion: Explorations in the Dynamics of Cognition.* Cambridge, MA: MIT Press.
van Fraassen, B. (1989). *Laws and Symmetry.* New York: Oxford University Press.
van Gelder, T. J. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences* 21: 1–14.
van Gelder, T., and R. Port. (1995). It's about time: An overview of the dynamical approach to cognition. In R. Port and T. van Gelder, Eds., *Mind as Motion: Explorations in the Dynamics of Cognition.* Cambridge, MA: MIT Press.
Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review* 20: 158–177.
Wilson, R. A. (1995). *Cartesian Psychology and Physical Minds: Individualism and the Sciences of the Mind.* New York: Cambridge University Press.
Wilson, R. A., Ed. (1999). *Species: New Interdisciplinary Essays.* Cambridge, MA: MIT Press.
Wundt, W. (1900–1909). *Völkerpsychologie.* Leipzig: W. Engelmann.
Yablo, S. (1992). Mental causation. *Philosophical Review* 101: 245–280.