20 *Philosophy of Mind: 1950–2000*

I want to outline some of the main developments in the philosophy of mind in the last half of the twentieth century.[1]

Behaviorism dominated psychology during approximately the same period that logical positivism dominated philosophy. The principles of behaviorism are less easily stated than those of logical positivism. It is perhaps better seen as a method that eschewed use of mentalistic vocabulary in favor of terms that made reference to dispositions to behavior. Both movements aimed at banishing nonscientific speculation, and forcing theory to hew as closely as possible to methods of confirmation. Both methodological doctrines came to be seen as restrictive, even on the practice of science.

Behaviorism had a run of influence within philosophy. It was a favored view of some of the later positivists. They made use of the verificationist principle to attempt to dissolve the mind–body problem and the problem of other minds, declaring these problems meaningless. And they appealed to behavioral analyses of mentalistic terms as a way of maintaining strict experimental control on mentalistic language. The simplistic picture of confirmation associated with the verificationist principle, a picture that ignored the role of auxiliary hypotheses, paralleled and abetted the behaviorist blindness to the role of background assumptions in mentalistic attributions. As we shall see, this blindness led to the collapse of behaviorism.

In postwar, postpositivistic philosophy, the early logical constructionists thought that behavioristic language was the most suitable way to 'reconstruct' mentalistic language in scientific terms. Ordinary-language philosophers purported to find behavioristic underpinnings for ordinary language. Behaviorism influenced positivistic construals of psychology, Quine's theory of the indeterminacy

[1] This article consists mostly of the second half of my article 'Philosophy of Language and Mind, 1950–1990', *The Philosophical Review*, 101 (1992), 3–51. I have added further material that concentrates on the last decade of the century. I present a historical overview pitched to nonspecialists. The scope of the article has, of course, led to omission of many important topics—for example, personal identity, action theory, the innateness of mental structures, knowledge of language, the nature of psychological explanation, the nature of concepts, many strands in the mind–body problem, and the legacy of Wittgenstein. I am grateful to Ned Block, Susan Carey, and the editors of *The Philosophical Review* for good advice.

of translation, Ryle's work on the concept of mind, and Malcolm's explications of discourse about dreaming and sensations.[2] These philosophers shared a tendency to think that theorizing in psychology or philosophy of mind should dispense with mentalistic vocabulary, or interpret it in nonmentalistic terms, as far as possible. They thought that such vocabulary should be largely replaced with talk about stimulations and about dispositions to behavior. Some philosophers thought that ordinary mentalistic terms could be defined or adequately explicated (for any cognitively respectable purpose) in these latter terms. Others thought that ordinary mentalistic terms were hopelessly unscientific or philosophically misleading, so no real explication was possible.

The demise of behaviorism in philosophy is less easily attributed to a few decisive events than is the fall of logical positivism. There were a series of influential criticisms of behaviorism beginning in the late 1950s and extending on for a decade.[3] The main cause of the shift seemed, however, to be a gradually developed sense that behaviorist methods were unduly restrictive and theoretically unfruitful. A similar development was unfolding within psychology, linguistics, and computer science, with an array of nonbehaviorist articles in the late 1950s and early 1960s.[4]

The attempts to provide behavioristic *explications* of mentalistic terms fell prey to various instances of a single problem. The behavioristic explications succeeded only on the implicit assumption that the individual had certain background beliefs or wants. As a crude illustration, consider an explication of belief as a disposition to assert. Even ignoring the fact that 'assert' is not a behavioral notion, but presupposes assumptions about mind and meaning, the analysis could work only with the proviso that the subject wants to express his beliefs and knows what they are. Eliminating these mentalistic background assumptions proved an impossible task, given behaviorist methodological strictures.

[2] Gilbert Ryle, *The Concept of Mind* (London: Hutchison, 1949); Norman Malcolm, *Dreaming* (London: Routledge & Kegan Paul, 1959); W. V. Quine, *Word and Object* (Cambridge, Mass.: MIT Press, 1960).

[3] Roderick Chisholm, *Perceiving* (Ithaca, NY: Cornell University Press, 1957), ch. 11; Peter Geach, *Mental Acts* (London: Routledge, 1957), ch. 1; Noam Chomsky, review of *Verbal Behavior*, by B. F. Skinner, *Language*, 35 (1959), 26–58, repr. in J. A. Fodor and J. Katz (eds.), *The Structure of Language* (Englewood Cliffs, NJ: Prentice-Hall, 1964) Hilary Putnam, 'Brains and Behavior' (1963), in *Philosophical Papers*, ii (Cambridge: Cambridge University Press, 1975); Jerry Fodor, *Psychological Explanation* (New York: Random House, 1968).

[4] In psychology: George Miller, 'The Magic Number 7 Plus or Minus Two: Some Limits on Our Capacity for Processing Information', *Psychological Review*, 63 (1956), 81–97; J. Bruner, J. Goodnow, and G. Austin, *A Study of Thinking* (New York: John Wiley, 1956); G. Miller, E. Galanter, and K. Pribram, *Plans and the Structure of Behavior* (New York: Holt, Rinehart & Winston, 1960); G. Sperling, 'The Information Available in Brief Visual Presentations', *Psychological Monographs*, 24 (1960); Ulrich Neisser, 'The Multiplicity of Thought', *British Journal of Psychology*, 54 (1963), 1–14; M. I. Posner, 'Immediate Memory in Sequential Tasks', *Psychology Bulletin*, 60 (1963), 333–349; S. Sternberg, 'High-Speed Scanning in Human Memory', *Science*, 153 (1966), 652–654. In linguistics: Noam Chomsky, *Syntactic Structures* (The Hague: Mouton, 1957). In computer science: A. Newell, J. C. Shaw, and H. A. Simon, 'Elements of a Theory of Human Problem Solving', *Psychological Review*, 65 (1958), 151–166.

The problem, stated less methodologically, is that mental causes typically have their behavioral effects only because of their interactions with one another.

As behaviorism slipped from prominence in philosophy in the 1950s and early 1960s, it left two heirs, which gradually formed an uneasy alliance. One of these heirs was naturalism. The other was functionalism.

A doctrine I will call 'naturalism' (and sometimes called 'physicalism') emerged first as a distinctive point of view in the philosophy of mind in the early 1950s. This view maintains two tenets. One is that there are no mental states, properties, events, objects, sensations over and above ordinary physical entities, entities identifiable in the physical sciences or entities that common sense would regard as physical. The formulation's vague expression 'over and above' matches the doctrine's vagueness: the doctrine does not entail an identity theory in ontology. It does require some sort of materialism about the mind. Naturalism coupled this ontological position with an ideological or methodological demand. It demanded that mentalistic discourse be reduced, explained, or eliminated in favor of discourse that is 'acceptable', or on some views already found, in the natural or physical sciences. Thus, we find repeated calls for 'explaining' rationality or intentionality. In its materialism, naturalism emphasized ontology in a way that behaviorism did not. Its ideological program, however, continued the behaviorist attempt to make psychology and philosophy of mind more scientific by limiting the supposed excesses of mentalism.

Many of the later logical positivists were naturalists. But issues about mind tended to be submerged in the general positivist program. The mind–body problem began to receive direct attention from a naturalistic point of view in articles by Quine, Place, and Smart, in the 1950s.[5] Place and Smart tried to identify mental states and events—primarily sensations and after images—with physical states and events. Smart thought that one could identify types of sensations in a 'topic-neutral' way that would leave it open whether they were physical; he then predicted that each type of sensation would turn out to be a neural state of some kind. For example, he paraphrased 'I am having an afterimage of an orange' as 'I am in a state like the one I am in when I am seeing an orange'. He thought that this translation would overcome any conceptual obstacles to identifying mental states with physical states. It would sidestep, for example, issues about the qualitative properties of afterimages. Science was supposed to settle the mind–body problem empirically—in favor of what came to be known as *type–type identity theory*, or *central state materialism*.

During the mid to late 1960s materialism became one of the few orthodoxies in American philosophy. It is difficult to say why this happened. No single

[5] W. V. Quine, 'On Mental Entities' (1952), in *The Ways of Paradox* (New York: Random House, 1966); U. T. Place, 'Is Consciousness a Brain Process?', *British Journal of Psychology*, 47 (1956), 44–50; J. J. C. Smart, 'Sensations and Brain Processes', *The Philosophical Review*, 68 (1959), 141–156.

argument obtained widespread acceptance. Perhaps the success in biochemistry during the 1950s in providing some sense of the chemical underpinnings of biological facts encouraged the expectation that eventually mental facts would receive a similar explication in neural terms. Moreover, there were some spectacular advances in animal neurophysiology during the period.[6] Perhaps the attempts of the positivists and behaviorists to make philosophy scientific had as a natural outgrowth the view that philosophical problems would eventually be solved by progress in the natural sciences—with the help of analytical clarification by philosophers. In any case, several philosophers in the 1960s defended either some form of the type–type identity theory or some form of eliminationism (the view that mentalistic talk and mental entities would eventually lose their place in our attempts to describe and explain the world).[7]

The most influential paper of this period was written several years before: Sellars's 'Empiricism and the Philosophy of Mind' (1956). The article is a grand attempt to portray mental episodes as explanatory posits that hold a place in our conceptual scheme by virtue of their explanatory usefulness.[8] Sellars tried to undermine the view that knowledge of one's own mental events is intrinsically poses an obstacle to the empirical discovery that mental events are neural events. Although in my view the argumentation in this paper is not satisfyingly clear or convincing, the picture it paints of the status of mentalistic discourse is profoundly conceived.

Whereas materialism became widely accepted during the 1960s, issues surrounding naturalism's ideological demand remained intensely controversial. Putnam raised a serious objection to type-type identity theories of the sort that Smart had made popular. He suggested that it is implausible that a sensation like pain is identical with a single neural state in all the many organisms that feel pain, in view of their enormously varied physiologies. He also pointed out that it is even more implausible to think that any given type of thought—for example, a thought that thrice 3 is 9 or a thought that one's present situation is dangerous—is realized by the same physical state in every being that thinks it. Not only the probable existence of extraterrestrials, the variety of higher

[6] J. Y. Lettvin *et al.*, 'What the Frog's Eye Tells the Frog's Brain', *Proceedings of the Institute of Radio Engineers*, 47 (1959), 1940–1951; D. H. Hubel and T. N. Wiesel, 'Receptive Fields of Single Neurones in the Cat's Striate Cortex', *Journal of Physiology*, 148 (1959), 574–591; Hubel and Wiesel, 'Receptive Fields, Binocular Interaction, and Functional Architecture in the Cat's Visual Cortex', *Journal of Physiology* (London), 160 (1962), 106–154.

[7] The central state identity theory is defended in D. M. Armstrong, *A Materialist Theory of the Mind* (London: Routledge & Kegan Paul, 1968); David Lewis, 'An Argument for the Identity Theory', *The Journal of Philosophy*, 63 (1966), 17–25. Eliminative materialism, which derives from Quine, is defended in Paul Feyerabend, 'Materialism and the Mind–Body Problem', *The Review of Metaphysics*, 17 (1963), 49–66; Richard Rorty, 'Mind–Body Identity, Privacy, and Categories', *The Review of Metaphysics*, 19 (1965), 24–54; and Daniel Dennett, *Content and Consciousness* (New York: Routledge & Kegan Paul, 1969). Many of these works, and several other significant ones, are collected in O'Connor (ed.), *Modern Materialism: Readings on Mind–Body Identity* (New York: Harcourt, Brace, and World, 1969).

[8] In Wilfrid Sellars, *Science, Perception, and Reality* (London: Routledge & Kegan Paul, 1963).

animals, and the possibility of thinking robots (a possibility most materialists were eager to defend), but the plasticity of the brain seemed to make the type–type identity theory untenable.[9] Mental states seemed 'multirealizable'. Materialism maintained its dominance, but needed a new form. Putnam observation seemed to show that if mentalistic discourse was to be explicated in 'scientifically acceptable' terms, the terms would have to be more abstract than neural terms.

Responses to Putnam's observation led to a more specific materialist orthodoxy. The response proceeded on two fronts: ontological and ideological. Most materialists gave up the type–type identity theory in favor of an ontology that came to be known as the token identity theory. Although a mental state- or event-kind was not identified with any one physical (neural) kind, each instance of a mental state and each particular mental event token was held to be identical with some instance of a physical state or with some physical event token. This claim allowed that the occurrence of a thought that thrice 3 is 9 could be identical with the occurrence of one sort of physical event in one person, whereas a different occurrence of the same kind of thought could be identical with the occurrence of a different sort of physical event in another person.

Although this ontological position is still widely maintained, no one argument for it has gained wide acceptance. The commonest consideration adduced in its favor is its supposed virtue in simplifying our understanding of mind–body causation. Davidson gave a profound but controversial apriori argument along these lines.[10] He held, first, that there are causal relations between mental and physical events; second, that causal relations between events must be backed by laws of a complete, closed system of explanation ('backed' in the sense that the predicates of the laws must be true of the events that are causally related); third, that there are no psycho-physical or purely mentalistic laws that form a complete, closed system of explanation. He concluded that since there can be no psycho-physical or mentalistic laws that would provide the relevant backing for the causal relations between mental and physical events, there must be purely physical laws that back such relations. This is to say that physical predicates apply to mental events—that mental events are physical.

Davidson has not been ideally clear or constant in formulating and arguing for the third premise. But given the conception of 'complete, closed system' that he usually adverts to, this premise seems plausible. The second premise is more doubtful. I do not think it apriori true, or even clearly a heuristic principle of science or reason, that causal relations must be backed by any particular kind

[9] Hilary Putnam, 'The Nature of Mental States' (1967), in *Philosophical Papers*, ii; Ned Block and Jerry Fodor 'What Psychological States Are Not', *The Philosophical Review*, 81 (1972), 159–181.

[10] Donald Davidson, 'Mental Events' (1970), in *Essays on Actions and Events* (Oxford: Clarendon Press, 1980).

of law. I think that we learn the nature and scope of laws (and the variety of sorts of 'laws') that back causal relations through empirical investigation. It is not clear that psycho-physical counterfactual generalizations—or nonstrict 'laws'—cannot alone 'back' psychophysical causal relations.

Most philosophers accepted the token identity theory as the simplest account that both reconciled materialism with multirealizability and raised no metaphysical issues about mind–body causation. Insofar as the view rests on the hope of finding empirical correlations between types that would inductively support token identities, however, it seems highly speculative. Some philosophers adopted an even more liberal materialism. They held, roughly, that although an instance of a mental event kind may not be an instance of a physical natural kind, they are always *constituted* of events that are instances of physical natural kinds.[11]

In any case, materialism in one form or another has widespread support among North American philosophers, largely on grounds of its supposed virtues in interpreting causation between mental and physical events. There is a vague sense abroad that alternatives amount to superstition. One common idea is that there is some intrinsic mystery in seeing mental events, imagined as nonphysical, as interacting with physical events. Descartes thought this too; and perhaps there was some plausibility to it, given his conceptions of mental and physical substance. But Cartesian conceptions of substance are not at issue nowadays, and the exact nature of the problem in its modern form needs clearer articulation than it is usually given.

A better-reasoned argument along these lines goes as follows. Macrophysical effects depend on prior macrophysical states or events according to approximately deterministic patterns described by physical laws. Mental causes often give rise to physical movements of human bodies. If such causation did not consist in physical processes, it would yield departures from the approximately deterministic patterns described by physical laws. It would interfere with, disrupt, alter, or otherwise 'make a difference' in the physical outcomes. But there is no reason to think that this occurs. Physical antecedent states seem to suffice for the physical effects. Appeal to mentalistic causation that does not consist in physical causation appears, on this reasoning, to invoke physically ungrounded causation that requires us to doubt the adequacy of current forms of physical explanation, even within the physical realm. Not surprisingly, such invocation is widely thought to be unattractive.

[11] Geoffrey Hellman and Frank Wilson Thompson, 'Physicalist Materialism', *Noûs*, 11 (1977), 309–345; Richard Boyd, 'Materialism without Reductionism: What Physicalism Does Not Entail', in Ned Block (ed.), *Readings in Philosophy of Psychology*, i (Cambridge, Mass.: Harvard University Press, 1980). Another source of reformulations of materialism has been the discussion of supervenience principles. Cf. Jaegwon Kim, 'Causality, Identity, and Supervenience in the Mind–Body Problem', *Midwest Studies in Philosophy*, 4 (1979), 31–50. It is worth noting, however, that supervenience of the mental on the physical does not entail materialism.

This reasoning—and other parallel arguments focusing on the effect of physical processes on mental states—has some force, perhaps enough to nourish materialism indefinitely. But I think that materialism merits more scepticism than it has received in North American philosophy during the last two decades. At any rate, the argument just outlined is not as forceful as it may appear.

Why should mental causes of physical effects interfere with the physical system if they do not *consist in* physical processes? Thinking that they must surely depends heavily on thinking of mental causes on a physical model—as providing an extra 'bump' or transfer of energy on the physical effect. In such a context, instances of 'overdetermination'—two causes having the same effect—must seem to be aberrations. But whether the physical model of mental causation is appropriate is part of what is at issue. Moreover, the sense in which mental causes must 'make a difference' if they do not consist in physical processes is in need of substantial clarification. There are many ways of specifying differences they do make that do not conflict with physical explanations.

It seems to me that we have substantial reason, just from considering mentalistic and physicalistic explanatory goals and practice—before ontology is even considered—to think that mentalistic and physicalistic accounts of causal processes will not interfere with one another. They appeal to common causes (in explaining the physiology and psychology of cognitive processes, for example) and common or at least constitutively related effects (in physiological and psychological explanations of an instance of a man's running to a store, for example). It seems to me perverse, independently of ontological considerations, to assume that these explanations might interfere with one another. They make too few assumptions about one another to allow such an assumption.

There are surely *some* systematic, even necessary, relations between mental events and underlying physical processes. It seems overwhelmingly plausible that mental events depend on physical events in some way or other. But constitution, identity, and physical composition are relations that have specific scientific uses in explaining relations between entities invoked in physical chemistry and biochemistry. These relations so far have no systematic use in nonmetaphysical, scientific theories bridging psychology and neurophysiology. They seem to me to be just one set of possibilities for accounting for relations between entities referred to in these very different explanatory enterprises. Where science does not make clear use of such relations, philosophy should postulate them with some diffidence.

The apparent fact that there are no gaps in physical chains of causation and that mental causes do not disrupt the physical system is perhaps ground for some sort of broad supervenience thesis—no changes in mental states without some sort of change in physical states. But the inference to materialism is, I think, a metaphysical speculation that has come, misleadingly, to seem a relatively obvious scientific-commonsensical bromide.

The issue of mind–body causation is extremely complex and subtle. In recent years, this issue has become an object of intense interest. Much of the discussion

concerns 'epiphenomenalism'.[12] The causal picture that motivates materialism is so firmly entrenched that many philosophers have come to worry that mental 'aspects' of events really do not 'make a difference': Maybe mental 'aspects' or properties are causally inert and just go along for a ride on physical properties of physical events, in something like the way that relations between phenotypal properties of parents and their offspring ride inertly and parasitically on underlying causal relations characterized by the genetic properties of parents and offspring. I think that these worries can be answered, even within a materialist framework. But I think that the very existence of the worries is the main point of philosophical interest. The worry about epiphenomenalism is, in my view, a sign that materialist theories have done a poor job of accounting for the relation between mind–body causal interaction and mentalistic explanation. They have done little to account for the fact that virtually all our knowledge and understanding of the nature and existence of mental causation derives from mentalistic explanations, not from nonintentional functionalist or neurological accounts.[13]

We determine the nature of the causation, and the sort of laws or law-like generalizations that accompany it, by scrutinizing actual explanations in psychology and ordinary discourse. If there turned out to be no clear sense in which mental events fell under predicates that are uncontroversially physical, then it would seem reasonable to count the mental events nonphysical. As far as I can see, there is no reason to be anything but relaxed in the face of this possibility. I see no powerful, clearly articulated reason for worrying about the existence of mind–body causation, or the gaplessness of chains of physical events, if this possibility were realized. What counts in supporting our belief in mind–body causation is the probity of mentalistic explanations. As long as they are informative and fruitful, we can assume that they are relating genuine events, whatever their metaphysical status.

Otherwise put: The theme in naturalism that deserves the status of orthodoxy is not its materialism and not its demand that mentalistic discourse be given some ideologically acceptable underpinning. It is its implicit insistence that one not countenance any form of explanation that will not stand the scrutiny of scientific and other well-established, pragmatically fruitful methods of communal check and testing. (More crudely, it is the opposition to miracles and to postulation of unverified interruptions in chains of causation among physical events.) But the relevant methods are to be drawn from reflection on what works in actual

[12] Cf., e.g., Jaegwon Kim, 'Epiphenomenal and Supervenient Causation', *Midwest Studies in Philosophy*, 9 (1984), 257–270; Ernest Sosa, 'Mind–Body Interaction and Supervenient Causation', ibid. 271–281; Ned Block, 'Can the Mind Change the World?', in G. Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam* (Cambridge: Cambridge University Press, 1990).

[13] The lack of attention to our source of knowledge of mental causation is one reason why there has recently been a small outpouring of worries among materialists that a form of epiphenomenalism—the view that mentalistic properties or descriptions are causally irrelevant—must be taken seriously.

explanatory practice, not from metaphysical or ideological restrictions on these practices. These points are subject to various interpretations. But I think that taking them seriously motivates less confidence in materialist metaphysics than is common in North American philosophy.

I have been discussing ontological responses to Putnam's observation that various kinds of physical states could be, and are, associated with mental states of a given type. The ideological response to Putnam's observation was the development of a new paradigm for indicating how mental states could be given identifications in nonmentalistic terms. Philosophers looked not to neurophysiology but to computer programming as a source of inspiration. Identifying a mental state with some sort of abstract state of a computer appeared to avoid the problems of identifying mental kinds with neural kinds. And unlike the nonreductive forms of token-identity materialism, it promised means of explaining mentalistic notions in other terms, or at least of supplementing and illuminating mentalistic explanation. Most philosophers found the terms of this supplementation compatible with materialism. This new account came to be known as *functionalism.*[14]

The guiding intuition of functionalism was that what entirely determines what kind of state or event a mental state or event is, is its place in a causal or functional network in the mental life of the individual. The original stimulus to this view was a proposed analogy between the mind and a computer program. To specify such a program, one needed to specify possible inputs into the system, the operations that would pass the machine from one state to another, the states that the machine would pass through, and the output of the machine, given each possible input and given the states it was already in. The machine might be either deterministic or probabilistic. On most versions of functionalism, the internal states were to be specified purely in terms of their 'place' in the system of input and output—in terms of the possible dependency relations they bore to other states and ultimately to input and output. Input and output were to be specified in nonintentional, nonmentalistic terms. Types of mental states and events were supposed to be determined entirely by the relations of functional dependency within the whole system of input and output.

The notion of determination is subject to three main interpretations. One, the least ambitious and least reductive, claims only that each mental kind supervenes on a place in the functional system, in the sense that the individual would be in a different kind of mental state if and only if he were not in the

[14] Cf. A. M. Turing, 'Computing Machinery and Intelligence', *Mind*, 59 (1950), 433–460. Turing's article provided an impetus and a vivid illustration of the computer paradigm, but it was itself an expression of behaviorism about the mind. The papers that inspired machine functionalism were Hilary Putnam's 'Minds and Machines' (1960), 'Robots: Machines or Artificially Created Life?' (1964), and 'The Mental Life of Some Machines' (1967), in *Philosophical Papers*, ii. Putnam states an explicitly functionalist view in 'The Nature of Mental States' (1967), but the idea is not far from the surface of his earlier papers. A type of functionalism less tied to computers was proposed in Lewis, 'An Argument for the Identity Theory' (1966), and Armstrong, *A Materialist Theory of the Mind.*

functional state corresponding to that kind. The other two purport to say what mental kinds 'consist in'. One version ('analytic functionalism') claims that a functionalist specification of such relations explicates the meaning of mentalistic terms. Another ('scientific functionalism') makes the lesser claim that such a specification gives the true essence of mental kinds, in something like the way that molecular constitution gives the true essence of a natural kind like *water*. Both of these latter two versions claim that functionalist discourse provides the 'real explanatory power' latent in mentalistic explanation.[15].

Analytic and scientific functionalism are clearly liberalized heirs to behaviorism. They share with behaviorism the insistence on nonintentional specifications of input (stimulus) and output (response), and the belief that mentalistic explanation is somehow deficient and needs a nonmentalistic underpinning. They also expand on the behaviorist idea that mental states are individuated partly in terms of their relations. Whereas behaviorists focused largely on relations to behavior, functionalists included relations to other mental states, and relations to stimulating input into the system. This is an insight already present in Frege, who claimed that sense is inseparable from a network of inferential capacities.

It has been common to combine functionalism with token-identity materialism. Functionalism was supposed to provide insight into the nature of mental kinds, whereas token-identity materialism provided insight into the nature of mental particulars—into the instantiation of the mental kinds in particular individuals. The computer analogy seemed compelling to many: mentalistic discourse was a sort of gloss on an underlying network functional flow chart, which was ultimately realized in different physical ways in different machines or organisms. Thus neural descriptions were seen as lying at the bottom of a three-level hierarchy of descriptions of the same human subject.

The functionalist position—in its least reductionist garb—was given distinctive form by Fodor. Fodor maintained that the intentional content of propositional attitudes is irreducible via functionalist specifications. But he held that such content is expressed by inner mental representations that have syntactic properties, inner words and sentences that were presumed to be instantiated somehow in the brain. Fodor further claimed that mental representations have their causal roles in virtue of their formal or syntactic properties, and that the input and output of functionalist specifications should be seen as symbols.[16] This picture

[15] The nonreductive version is the least common. It is expressed in the introduction of Jerry Fodor's *RePresentations* (Cambridge, Mass.: MIT Press, 1981), but he maintains it neither very long before nor very long after. The analytic version may be found in Armstrong, *A Materialist Theory of Mind*; David Lewis, 'Psychophysical and Theoretical Identification', *Australasian Journal of Philosophy*, 50 (1972), 249–258; Sydney Shoemaker, 'Functionalism and Qualia' (1975), in his *Identity, Cause and Mind* (Cambridge: Cambridge University Press, 1984). Putnam proposed the scientific version in 'The Nature of Mental States'. A view more instrumentalist than functionalist but which bears broad comparison appears in Daniel Dennett, 'Intentional Systems', *The Journal of Philosophy*, 68 (1971), 87–106.

[16] Jerry A. Fodor, *The Language of Thought* (New York: Cravell, 1975) and *RePresentations*. Cf. also Hartry Field, 'Mental Representation', *Erkenntnis*, 13 (1978), 9–61.

brought the functionalist tradition into line with a fairly literal interpretation of the computer analogy: psychological explanation was modeled on *proofs* or other types of symbol manipulation by a digital computer. The causal aspects of psychological explanation were to be understood in terms of the physical relations among the particular neural states or events that instantiated the symbolic representations.

Something like this picture had been proposed by Sellars.[17] But Fodor presented his view as an interpretation of work in psycholinguistics and cognitive psychology. To many it gained plausibility because of its appeal to specific scientific practices. The picture and its relation to psychological theory are still very much in dispute.[18] Fodor's work drew attention from linguists, psychologists, and computer scientists. It also benefited from and helped further a significant shift in the degree to which the details of scientific practice were seen to be relevant to philosophical problems about mind.

Until the mid to late 1970s most philosophy in this area was carried on in a relatively apriori analytic spirit. Even those philosophers, such as type–type identity theorists or sceptics about mental states, who purported to take science as a model for philosophy of mind had little to say about the theories of any science. They saw themselves as freeing philosophy from obstacles to scientific progress (whose direction was often predicted with considerable confidence). This was true not only of the philosophy of mind, but of much of the rest of philosophy—even much of the philosophy of natural science, with the exception of historical work in the tradition of Thomas Kuhn.[19] It is an interesting question why such a shift occurred. A similar shift occurred in the philosophies of science and mathematics. Both disciplines undertook much more concentrated discussions of a wider variety of the details of scientific practice, beginning about fifteen years ago.[20] Philosophizing about biology, a science that had not

[17] Wilfrid Sellars, 'Some Reflections on Language Games' (1954), in *Science, Perception and Reality*. Cf. also Gilbert Harman, *Thought* (Princeton: Princeton University Press, 1973).

[18] For opposition from different angles to the computer analogy or to other aspects of the language-of-thought hypothesis, see Paul M. Churchland, *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge University Press, 1979); Christopher Peacocke, *Sense and Content* (Oxford: Clarendon Press, 1983); Stephen Stich, *From Folk Psychology to Cognitive Science* (Cambridge, Mass.: MIT Press, 1983); Robert Stalnaker, *Inquiry* (Cambridge, Mass.: MIT Press, 1984); Daniel Dennett, *The Intentional Stance* (Cambridge, Mass.: MIT Press, 1987); Paul Smolensky, 'On the Proper Treatment of Connectionism', *The Journal of Behavioral and Brain Sciences*, 11 (1988), 1–74.

[19] T. S. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962).

[20] The change in the philosophy of physics was foreshadowed by early articles of Hilary Putnam's—e.g. 'An Examination of Grünbaum's Philosophy of Geometry' (1963), 'A Philosopher Looks at Quantum Mechanics' (1965), both in *Philosophical Papers*, i (Cambridge: Cambridge University Press, 1975). But it caught on and received new impetus with the articles of John Earman—e.g. 'Who's Afraid of Absolute Space?', *Australasian Journal of Philosophy*, 48 (1970), 287–319. For an overview of broadly analogous changes in the philosophy of mathematics, see Thomas Tymoczko (ed.), *New Directions in the Philosophy of Mathematics* (Boston: Birkhauser, 1985).

conformed to positivist conceptions of law and explanation, came to prominence in this period.

Perhaps it took two decades for the criticisms of positivism to be digested sufficiently for a more open-minded consideration of the actual practice of the sciences to develop. In any case, interest in the details of psychology should be seen in the context of intellectual movements outside the scope of this essay.

The demise of behaviorism might similarly be viewed as requiring a period of assimilation before psychology could be considered a worthwhile object of philosophical reflection. Of course, there was a more positive side to the reconsideration of the practice of psychology. The computer paradigm was a natural object of interest. The continuing success of Chomsky's program in linguistics, coupled as it was with claims that it was a part of a psychology of the mind, made philosophers increasingly interested in mentalistic psychology. And an intellectually substantial cognitive and developmental psychology, and psycholinguistics, offered new forms to questions relevant to traditional philosophical issues: the role of intentional content in explanation, the mind–body problem, differences between the natural and the human sciences, the relation between language and thought, the innateness and universality of various conceptual and linguistic structures, the scope and limits of human rationality.

How much the reflection on psychology will enrich and advance philosophical inquiry remains an open question. Quite a lot of the work in this area seems to me very unreflective. It is at best rare that scientific practice answers philosophical questions in a straightforward way. But philosophy has traditionally given and received aid in the rise of new sciences or new scientific paradigms.

Let us return to functionalism. Although functionalism has enjoyed substantial support—at least among specialists in the philosophy of mind—it has not lacked detractors. The analytic and scientific versions of functionalism have always been afflicted with a programmatic, unspecific character that has seemed to many to render them unilluminating as *accounts* of particular mental kinds.

There are more specific criticisms. Many philosophers find the application of any form of functionalism to sensations like pain or color sensations implausible. For them, the causal relations of the sensations seem less fundamental to their character than their qualitative aspects.[21]

Searle mounted a controversial argument, similar to some of those directed against the applicability of functionalism to qualitative aspects of sensations,

[21] Criticism of this aspect of functionalism may be found in Ned Block, 'Troubles with Functionalism', in C. W. Savage (ed.), *Minnesota Studies in the Philosophy of Science*, ix (Minneapolis: University of Minnesota Press, 1978), and 'Are Absent Qualia Impossible?' *The Philosophical Review*, 89 (1980), 257–274. An influential article with a different, but related, point is Thomas Nagel, 'What Is It Like to Be a Bat?', *The Philosophical Review*, 83 (1974), 435–450. Cf. also Frank Jackson, 'Epiphenomenal Qualia', *Philosophical Quarterly*, 32 (1982), 127–136. The numerous defenses of functionalism on this score include Sydney Shoemaker, 'Functionalism and Qualia' and 'Absent Qualia are Impossible—A Reply to Block', in *Identity, Cause and Mind*; and David Lewis, 'Mad Pain and Martian Pain' (1980), in his *Philosophical Papers*, i (New York: Oxford University Press, 1983).

to show that functionalism could not account for any propositional attitudes. He postulated a room in which stations are manned by a person who does not understand Chinese, but who memorizes the Chinese words of given instructions. These stations are postulated to correspond to the stages of processing a language. The person is able to produce appropriate Chinese sentences as output, given any Chinese sentence as input. Searle claimed that although the system could be set up to meet the functionalist requirements for understanding Chinese, there is no understanding of Chinese in the room. Most opponents claim that the whole system can be credited with understanding Chinese. Searle finds this reply unconvincing.[22]

A more complex issue concerns the specific formulation of a functionalist account. Clearly, people can share meanings and many beliefs even though they maintain very different theories about the world. Maintaining different theories entails making different inferences, which correspond to different causal relations among the different sets of mental states associated with the theories. So not just any network of causal relations among mental states and events can be relevant to a functional account, on pain of counting no one as sharing any beliefs or meanings. One needs to find a network that is common to all the possible inference networks and theories in which any given belief (or meaning) might be embedded. But it is very difficult to imagine there being such common causal networks for each given belief (or meaning).[23]

Another approach to understanding intentional content and mental kinds developed out of the work on reference. That work showed that proper names and natural kind expressions could succeed in referring even though the speaker's knowledge of the referent was incomplete or defective. Reference depends not just on background descriptions that the speaker associates with the relevant words, but on contextual, not purely cognitive relations that the speaker bears to entities that a term applies to.

The work on reference is relevant to the meaning of terms and to the identity of concepts. For the meaning of a wide range of nonindexical terms and the nature of a wide range of concepts are dependent on the referent or range of application in the sense that if the referent were different, the meaning of the term, and the associated concept, would be different. (Here let us simply take concepts to be elements in the intentional contents of propositional attitudes, elements that have referential aspects.) For example, different meanings or concepts would be expressed by the wordforms 'chair' and 'arthritis' if the word forms did not apply exactly to chairs and to instances of arthritis.

[22] John Searle, 'Minds, Brains, and Programs', *The Behavioral and Brain Sciences*, 3 (1980), 417–424. Searle's argument is anticipated in Ned Block, 'Troubles with Functionalism'.

[23] These problems have long been recognized. But as with some of the fundamental difficulties with positivism, such recognition does not always convince proponents of a program to give it up. For a summary of some of these problems, see Hilary Putnam, *Representation and Reality* (Cambridge, Mass.: MIT Press, 1988).

The points about reference can be extended to many such terms and concepts. An individual can think of a range of entities via such terms and concepts even though the thinker's knowledge of the entities is not complete enough to pick out that range of entities except through the employment of those terms and concepts. What the individual knows about the range of entities—and hence about those many meanings or concepts whose identities are not independent of their referential range of applications—need not provide a definition that distinguishes them from all other (possible) meanings or concepts. So the meanings of many terms—and the identities of many concepts—are what they are even though what the individual knows about the meaning or concept may be insufficient to determine it uniquely. Their identities are fixed by environmental factors that are not entirely captured in the explicatory or even discriminatory abilities of the individual, unless those discriminatory abilities include application of the concept itself. Since most propositional attitudes, like specific beliefs, are the kinds of mental kinds that they are because of the meanings, concepts, or intentional contents that are used to specify them, the identities of many mental kinds depend on environmental factors that are not entirely captured in the (nonintentionally specified) discriminatory abilities of the individual. I have just developed one motivation for what is called '*anti-individualism*'.

Anti-individualism is the view that not all of an individual's mental states and events can be type-individuated independently of the nature of the entities in the individual's environment. There is, on this view, a deep individuative relation between the individual's being in mental states of certain kinds and the nature of the individual's physical or social environments.

Anti-individualism was supported not only through abstract considerations from the theory of reference, but also through specific thought experiments. For example, one can imagine two individuals who are, for all relevant purposes, identical in the intrinsic physical nature and history of their bodies (described in isolation of their environments). But the two individuals can be imagined to have interacted with different metals (one aluminum, one an aluminum look-alike) in their respective environments. The metals need resemble one another only to the level of detail that the two individuals have noticed. The individuals know about as much about the metals as most ordinary people do, but neither could tell the difference if given the other metal. In such a case, it seems that one individual has thoughts like *aluminum is a light metal*, whereas the other individual (lacking any access to aluminum, even through interlocutors) has analogous thoughts about the other metal. Similar thought experiments appear to show that a person's thoughts can be dependent on relations to a social environment as well as a purely physical one. Some environmental dependence or other can be shown for nearly all empirically applicable terms or concepts.[24]

[24] Tyler Burge, 'Individualism and the Mental', *Midwest Studies in Philosophy*, 4 (1979), 73–121; 'Other Bodies', in A. Woodfield (ed.), *Thought and Object* (Oxford: Oxford University Press, 1982); 'Intellectual Norms and Foundations of Mind', *The Journal of Philosophy*, 83 (1986),

The thought experiments made trouble for the standard forms of functionalism, which limited specifications of input and output to the surfaces of the individual. The thought experiments suggested that all an individual's internal functional transactions could remain constant, while his mental states (counterfactually) varied. Some philosophers proposed extending the functional network into the physical or social environments. Such a proposal reduces the reliance on the computer paradigm and requires a vastly more complex account. The main problems for it are those of accounting for (or specifying an illuminating supervenience base for) the notions of meaning, reference, and social dependence, in nonintentional terms. These are tasks commonly underestimated, in my view, because of the programmatic nature of the functionalist proposals.

Most philosophers seem to have accepted the thought experiments. But there remains disagreement about how they bear on mentalistic explanation, especially in psychology. Some have held that no notion of intentional content that is as dependent for its individuation on matters external to the individual could serve in explaining the individual's behavior. Many of these philosophers have tried to fashion surrogate notions of content or of 'mental' states to serve explanatory purposes. Others have maintained that such positions are based on mistakes and that the ordinary notions of intentional content and mental state can and do play a role in ordinary explanation and explanation in psychology. The debate concerns the interpretation of actual psychological practice and the relation between psychological explanation and explanation in other sciences.[25]

In my view, however, the main interest of the thought experiments lies in their giving new forms to many old issues. The arguments for anti-individualism are new. But the broad outline of the conclusion that they support is not. It is

697–720; 'Cartesian Error and the Objectivity of Perception', in R. Grimm and D. Merrill (eds.), *Contents of Thought* (Tucson: University of Arizona Press, 1988); 'Wherein is Language Social?', in A. George (ed.), *Reflections on Chomsky* (Oxford: Basil Blackwell, 1989) (Chs. 5, 4, 10, 7, 11 above). The thought experiments use the methodology set out in Hilary Putnam, 'The Meaning of "Meaning"' (1975), in *Philosophical Papers*, ii. Putnam's argument, however, was not applied to intentional elements in mind or meaning. In fact, it contained remarks that are incompatible with anti-individualism about mental states. Much in subsequent papers is, however, anti-individualistic. Cf. 'Computational Psychology and Interpretation Theory', in *Philosophical Papers*, iii (Cambridge: Cambridge University Press, 1983); *Representation and Reality*, ch. 5. But ambivalences remain. Cf. ibid, 19–22.

[25] For versions of the former approach, see Stephen White, 'Partial Character and the Language of Thought', *Pacific Philosophical Quarterly*, 63 (1982), 347–365; Stephen Stich, 'On the Ascription of Content', in *Thought and Object*; Jerry Fodor, *Psychosemantics* (Cambridge, Mass.: MIT Press, 1987); Brian Loar, 'Social Content and Psychological Content', in Grimm and Merrill (eds.), *Contents of Thought*. For defenses of anti-individualistic conceptions of psychology, see Fred Dretske, *Knowledge and the Flow of Information* (Cambridge, Mass.: MIT Press, 1981); Tyler Burge, 'Individualism and Psychology', *The Philosophical Review*, 95 (1986), 3–45 (Ch. 9 above), and 'Individuation and Causation in Psychology', *Pacific Philosophical Quarterly*, 70 (1989), 303–22 (Ch. 14 above); Lynne Rudder Baker, *Saving Belief* (Princeton: Princeton University Press, 1987); and Robert Stalnaker, 'On What's in the Head', *Philosophical Perspectives*, 8 (1989), 287–316.

clearly maintained by Aristotle, Hegel, and Wittgenstein, and arguably present in Descartes and Kant.[26] Emergence of an old doctrine in a new form is a source of vitality in philosophy. Issues about self-knowledge, skepticism, apriori knowledge, personhood, the nature of meaning, the mind-body problem, are all deeply affected by considerations about necessary, individuative relations between an individual's mind and his environment. The line of development from the anti-descriptivist theories of reference to anti-individualist accounts of mind promises, I think, to enrich traditional philosophy.

In the last decade of the twentieth century, the relation between anti-individualism and other issues in philosophy came under intense scrutiny. I will discuss two areas that fall under this general heading. One is perception and perceptual thought. The other is self-knowledge. The decade was also marked by the emergence of widespread reflection on qualitative aspects of the mental, and on the nature of consciousness.

I begin with the two issues associated with anti-individualism. The first concerns issues about singular *de re* aspects of representation. I touch on two issues within this sub-area, both having to do with the nature of perceptual representation.

In *Languages of Art*, Nelson Goodman developed an account of the 'syntax' of pictorial representation that distinguishes it from propositional representation. Many have thought that Goodman's work points toward an account of nonpropositional form for perceptual representation. Roughly, Goodman counted nonpropositional representations—particularly drawings—as analog representations. Analog representations are analog if (for relevant purposes) dense. Representations are dense if between any two types there is a third. An associated idea is that in analog representations, every discernible difference in the representational medium makes a representational difference. There is much that is not right about Goodman's account, even for pictures. Still, many have thought that Goodman was on to an important distinction between perceptual and conceptual (propositional) representation, particularly regarding the second idea. Others pursued a similar path independently.[27]

[26] Descartes's Demon hypothesis is paradigmatically individualistic. But Descartes thought that the hypothesis was incoherent. His causal argument for the existence of the physical world (in *Meditation* VI) and his principle that the reality of ideas cannot exceed the reality of their objects are anti-individualistic in spirit. The question of whether Descartes was an individualist is very complex and entangled with his views about God. As regards Kant, the Refutation of Idealism (*Critique of Pure Reason*, B 274 ff.) contains a fundamentally anti-individualistic strategy. But the overall question of how to interpret Kant with regard to anti-individualism is, again, very complex, since it is bound up with the interpretation of his transcendental idealism.

[27] Nelson Goodman, *Languages of Art* (Indianapolis: Bobbs-Merrill, 1968); John Haugeland, 'Analog and Analog', *Philosophical Topics*, 12 (1981), 213–226; Fred Dretske, *Knowledge and the Flow of Information* (Oxford: Blackwell, 1981); Gareth Evans, *The Varieties of Reference* (Oxford: Oxford University Press, 1982); Christopher Peacocke, 'Perceptual Content', in J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan* (Oxford: Oxford University Press, 1989); *idem, A Study of Concepts* (Cambridge, Mass.: MIT Press, 1992).

Making the distinction in a clear and psychologically relevant way resists simple stories. There is the following baseline point that functions as a challenge: seemingly any representational content can be mimicked by or converted into propositional form. What distinguishes (some particular form of) nonpropositional, nonconceptual representational content? Since perception seems both to indicate particulars and to categorize them in certain ways, one has a prima facie analog of subject–predicate form implicit in perception.

Some have claimed that one cannot make sense of representation that plays a role in epistemology unless one takes the representation to be propositional and thus capable of yielding reasons.[28]

I believe that the position that distinguishes perceptual from conceptual representational form is correct. I believe this largely on empirical grounds. There is no need to appeal to propositionally marked states to explain the representational capacities, principally perceptual capacities, of numerous animals. But perception and action in these animals *is* best explained in terms of representational states. The representational organization of vision, hearing, and touch does not seem to be propositional. In most empirical work on these senses, no systematic attribution of propositional representations is made. Similarly, the representational organization of grasping or of eating does not seem to be propositional. No reasonable, full account can carry out the inquiry independently of empirical work in psychology. But the need for conceptual clarification in understanding the distinction seems to me to be deep and complex.

A related issue about perceptual representation concerns its semantics rather than its form. The issue centers on the nature of the relation between the singular element in perceptual belief and the particulars that the perceptual belief is about. Developing the view in anti-individualism that representational states are (commonly) individuated in terms of their relations to the environment, some, following Evans, maintained that a perceptual state and a perceptual belief could not be the same type of state or belief if it were a perception or belief about a different particular. Similarly, a given type of perceptual state or belief that is about a given particular could not have been the same if there had been an illusion of a particular instead of a genuine perceptual object in the environment.[29] On the other hand, there is both commonsense and scientific ground for thinking that the types of states that perceivers are in are the same as between veridical perceptions, perceptions of indiscernible duplicates, and

[28] John McDowell, *Mind and World* (Cambridge, Mass.: Harvard University Press, 1994). Cf. also the exchange between Peacocke and McDowell, *Philosophy and Phenomenological Research*, 57 (1998), 381–388, 414–419.

[29] Cf. Evans, *The Varieties of Reference*; John McDowell, 'Singular Thought and the Boundaries of Inner Space', in J. McDowell and P. Pettit (eds.), *Subject, Thought, and Context* (Oxford: Oxford University Press, 1986); Peacocke, *A Study of Concepts*; McDowell, *Mind and World*.

referential perceptual illusions.[30] This issue raises important questions about the nature of perceptual states and perceptual beliefs and about the nature of illusion.

A second large sub-area in the development of issues associated with anti-individualism bears on the nature of self-knowledge. Discussion centered on the question of whether anti-individualism is compatible with some sort of authoritative or privileged warrant for certain types of self-knowledge. Inevitably, this question forced reflection on the nature of self-knowledge and its role in various human pursuits. The issue was first raised in independent papers by Davidson and me. Each defended a type of compatibilism about the relation between anti-individualist individuation of certain mental states and a capacity to know nonempirically what those states are.[31] This claim was resisted or qualified by a number of philosophers.[32]

The bulk of the discussion came to center on the nature of self-knowledge—or rather, the nature of various types of self-knowledge. Some philosophers maintained that all self-knowledge is at least implicitly empirical. Some maintained that apparent cases of privileged or authoritative self-knowledge are to be understood in an expressivist or otherwise deflationary way. Others attempted to understand the apparently special character of some knowledge of what one is thinking or what one believes—without explaining that character away. What I regard as the most interesting developments of this approach appealed to some constitutive role of self-knowledge in having beliefs, in having a concept of belief, in being rational, or in being a deliberative person. These matters are complex, and deserve, I think, further development.[33]

I want to highlight one further area of intense discussion in the philosophy of mind in the 1990s. This one is largely independent of issues about anti-individualism. The area concerns the nature of qualitative experience, including

[30] John Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983); Tyler Burge, 'Vision and Intentional Content', in E. Lepore and R. Van Gulick (eds.), *John Searle and his Critics* (Oxford: Blackwell, 1991).

[31] Donald Davidson, 'Knowing One's Own Mind', *Proceedings and Addresses of the American Philosophical Association*, 60 (1987), 441–458; Tyler Burge, 'Individualism and Self-Knowledge', *The Journal of Philosophy*, 85 (1988), 649–663.

[32] Paul Boghossian, 'Content and Self-Knowledge', *Philosophical Topics*, 17 (1989), 5–26; André Gallois, *The World Without, the Mind Within* (Cambridge: Cambridge University Press, 1996). For a collection of articles developing both sides of this issue, see P. Ludlow and N. Martin (eds.), *Externalism and Self-Knowledge* (Stanford, Calif.: CSLI Publications, 1998).

[33] Donald Davidson, 'First Person Authority', *Dialectica*, 38 (1984), 101–110, repr. in *Subjective, Intersubjective, Objective* (Oxford: Oxford University Press, 2001); Sydney Shoemaker, 'Self-Knowledge and "Inner Sense" ', *Philosophy and Phenomenological Research*, 54 (1994), 249–314, repr. in *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press, 1996); Richard Moran, 'Interpretation Theory and the First-Person', *Philosophical Quarterly*, 44 (1994), 154–173; Tyler Burge, 'Our Entitlement to Self-Knowledge', *Proceedings of the Aristotelian Society*, 96 (1995), 1–26; Bernard Kobes, 'Mental Content and Hot Self-Knowledge', *Philosophical Topics*, 24 (1996), 71–99. A collection of articles that provides some indication of the range of this discussion is C. Wright, B. C. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds* (Oxford: Oxford University Press, 1998).

the nature of consciousness. Consciousness and representationality, or aboutness, have long been regarded as the two major marks of mind. An important question is whether one mark can be reduced to the other; and if not, what relative places the two marks have in our understanding of mind.

The discussion of these matters at the end of the twentieth century must be seen against the background of four important papers. I have already mentioned two of these: Block's 'Troubles with Functionalism'[34] and Searle's 'Minds, Brains, and Programs'.[35] Each paper offers a forceful example that suggests that functionalist accounts of the representational aspects of mind fail to come to grips with qualitative aspects of experience. A third paper, Nagel's 'What is it Like to Be a Bat?',[36] offers a compelling way of thinking about qualitative aspects of experience, summarized in his phrase 'what it is like'. Finally, Jackson's 'Epiphenomenal Qualia',[37] gives an argument against materialism that features the difficulty of accounting for phenomenal qualities in material terms. These papers set many cross-currents going. I will not be able to survey nearly all of them. I will concentrate on one strand of development—the relation between qualitative and representational aspects of mind.

Many of the original responses to these papers focused on defending either functionalism or materialism—the original targets of the papers. As interest in strict forms of functionalism waned, the debate over functionalism was largely replaced by a closely related, but slightly different debate. Some functionalists took the tack of *assuming* that some form of functionalism is true of representational states, and then arguing that qualitative phenomena are essentially and solely representational. Others argued to the same conclusion with no antecedent commitment to functionalism.

Harman's 'The Intrinsic Quality of Experience'[38] claimed that qualitative aspects of experience are simply certain types of representational aspects of experience. Qualitative aspects of experience are the aspects that have to do with what it is like to have an experience, feeling, or sensation. Harman argued that putatively intrinsic aspects of experiences—for example, the felt quality of pain—had been conflated with intrinsic aspects of the 'intentional object' of an experience. Against the Jackson thought experiment he maintained that a person blind from birth fails to know what it is like to see something red because he or she does not have the full concept of red and so does not fully understand what it is for something to be red. Finally, Harman argued against invoking the inverted spectrum to show that representational constancy is compatible with qualitative difference. Harman's paper defended what came to be known

[34] In C. W. Savage (ed.), *Minnesota Studies in the Philosophy of Science*, ix (Minneapolis: University of Minnesota Press, 1978).

[35] *The Behavioral and Brain Sciences*, 3 (1980), 417–424.

[36] *The Philosophical Review*, 83 (1974), 435–450.

[37] *Philosophical Quarterly*, 37 (1982), 127–136.

[38] In J. Tomberlin (ed.), *Philosophical Perspectives*, iv (Atascadero, Calif.: Ridgeview Publishing Co., 1990).

as 'representationalism'—the view that qualitative aspects of experience are nothing other than representational aspects.

This paper was followed by—and in many cases it engendered—several further papers defending representationalist construals of qualitative mental phenomena.[39] The position was opposed by other philosophers, who maintained that phenomenal qualities, or phenomenal qualitative aspects of experience, commonly have representational content and representational functions but are not to be reduced to them.[40]

Most papers in this area, on both sides, pay what seems to me too little attention to what is to be meant by 'representation'. Harman admitted that his notion of intentional object is crude. I think it vulnerable. I think that the notion will not convincingly support the first of Harman's three arguments. There is, however, a clearer conception of representation that provides a basis for much of the representationalist discussion. According to this conception '*x* represents *y*' should be understood roughly as: appropriate types of which *x* is an instance are dependent in a lawful or law-like way on appropriate types of which *y* is an instance, and this connection between *x* and *y* has functional value for the life of the organism that contains x.[41]

It is certainly plausible that on this conception of representation, qualitative color registrations are representational, if not constitutively, at least as a matter of fact. The view also suggests that pains and orgasms are representational. For example, a state of feeling pain is in a law-like relation to instances of bodily damage or disorder, and this relation surely has a functional value in the life of the organism. So the state of feeling pain 'represents' the bodily damage or disorder. This result has been embraced and defended by some representationalists.[42]

I believe that this notion of representation is too broad. It deems the states of very simple creatures to be 'representational states'. For example, it counts as representational states the sensory states underlying thermotactic responses of protozoa. One can coherently talk this way. Some psychologists and physiologists do talk this way. This notion of representation is, however, not needed to

[39] Fred Dretske, *Naturalizing the Mind* (Cambridge, Mass.: MIT Press, 1995); *idem.*, 'Phenomenal Externalism, or If Meanings Ain't in the Head, Where are Qualia?', in E. Villanueva (ed.), *Philosophical Issues, vii: Perception* (Atascadero, Calif.: Ridgeview, 1993); Michael Tye, *Ten Problems of Consciousness* (Cambridge, Mass.: MIT Press, 1995); Georges Rey, 'Sensational Sentences', in M. Davies and G. Humphreys (eds.), *Consciousness: Psychological and Philosophical Essays* (Oxford: Blackwell, 1992).

[40] Ned Block, 'Inverted Earth', *Philosophical Perspectives*, 4 (1990), 51–79; *idem.* 'Mental Paint and Mental Latex', in Villanueva (ed.), *Philosophical Issues, vii: Perception;*, Brian Loar, 'Phenomenal States', in N. Block, O. Flanagan, and G. Güzeldere (eds.), *The Nature of Consciousness* (Cambridge, Mass.: MIT Press, 1997; revision of original 1990 version); Colin McGinn, *The Problem of Consciousness* (Oxford: Blackwell, 1991).

[41] Dretske, *Naturalizing the Mind*, ch. 1. A similar but somewhat different view can be found in Ruth Millikan, *Language, Thought, and Other Biological Categories* (Cambridge, Mass.: MIT Press, 1984; and idem., 'Biosemantics', *The Journal of Philosophy*, 86 (1989), 281–297.

[42] Michael Tye, 'A Representational Theory of Pains and their Phenomenal Character', in J. Tomberlin (ed.), *Philosophical Perspectives*, ix (Atascadero, Calif.: Ridgeview, 1990).

explain such simple phenomena. It is a mere gloss on points that can be made in physiological, ecological, and functional terms. By contrast, a more intuitive, restrictive notion of representation does have prima facie explanatory force in the ethology of the perceptual and cognitive systems of less primitive animals and in the psychology of human beings. Dretske and others hope to reduce the more ordinary notion(s) of representation to the very broad one. I think that there is no ground for optimism about this project.[43]

Moreover, it is doubtful that the simple notion is one that opponents of representationalism have been concerned with when they doubt that qualitative aspects of experience are 'representational'. I believe that a less inclusive, or more specific, notion is needed to sort out issues between representationalists and their opponents.

Laying this issue aside, there remain difficulties for the representationalist position, even formulated with the broad notion of representation. One is that it produces unattractive results in cases like the feeling of pain. For example, one representationalist position holds that if an individual has a phantom limb, there is no pain—since there is no bodily damage that is functionally connected to the state of feeling pain; there is only a representation of pain that hallucinates the pain. This view has actually been embraced by some proponents of representationalism. It seems to me clearly unacceptable.

A second difficulty is that despite the inclusiveness of the cited conception of representation, there still appear to be qualitative states that do not 'represent' in this inclusive sense. There appear to be qualitative aspects of experience that have no function in the life of the organism. They constitute dysfunction or noise. Blurriness in a visual experience is an example. These cases have been treated as misrepresentations or as representations of blurriness, but it is hard to see that they have any representational function at all. They make no contribution to reproductive fitness, and they seem to get in the way of proper functions. This result is incompatible with the representationalist program.

Harman's second argument seems to me to beg the question. It does, however, highlight interesting and difficult issues about the relation between qualitative aspects of perception and representational aspects. In the case of (say) visual perception it is certainly hard to separate out qualitative aspects from representational aspects. We identify most qualitative aspects of perception with terms that indicate what is normally represented when one's consciousness is marked by those aspects. For example, although few philosophers still think that experience (by a normal-sighted person) of red is itself in any way red, we currently have no better, easily accessible, public way to characterize the qualitative aspect of the experience than in terms of its relation to red surfaces. This fact makes it

[43] Dretske relies partly on a view that in Mental representation, there is learning. In the relevant sense of learning, all animals and all unicellular organisms, including amoebae, learn. No animal is confined to fixed hard-wired, purely reflexive behavior. So Dretske's line seems to me not to do anything to solve the problem.

hard to see how to describe, much less explain, the qualitative aspect of experience in a way that abstracts from its representational role. And this difficulty encourages representationalism.

Opponents of representationalism typically maintain that the qualitative aspects of perception play a role in determining the mode of presentation or the representational content of the perception, and are not to be explained in terms of this role. The idea is that at least some aspects of qualitative states are dependent on the nature of the neural substrate and not on the sorts of correlations with objects of representation that determine representational content. This view seems to me correct. But it will provide little explanatory illumination until the relation between qualitative feels and neural underpinnings are better and more specifically understood.

Harman's third argument centers on the inverted spectrum. How to describe and evaluate this type of thought experiment remains difficult and disputed. I think that the case can be elaborated in various ways that make representationalism implausible. But the issues are very delicate and complex, and I will not try to review them here. There has been considerable, nuanced discussion of issues surrounding inverted spectra.[44]

One reason why the issue of representationalism has been of such interest is that it bears on the nature of consciousness. Qualitative or phenomenal states, as ordinarily felt, provide the most obvious instances of consciousness. Feelings involved in sensations like pains and tickles, phenomenal aspects of the experience of color or sound, the experience of warmth or of being touched, seem to be the paradigmatic center of consciousness. There are then issues about the relation between consciousness and representation that parallel those about the relation between qualitative aspects of experience and representation. In fact, positions on consciousness run the gamut from arguments that there is no such thing as qualitative aspects of experience, to functionalist reductionism, to simple representationalist views, to higher-order thought views, to claims that consciousness is irreducible, to claims that consciousness is both irreducible and a ground for dualist theories of mind.[45]

[44] Sydney Shoemaker, 'The Inverted Spectrum', *The Journal of Philosophy*, 74 (1982), 357–381; 'Intrasubjective/Intersubjective', in *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press, 1996); Block, 'Inverted Earth'.

[45] Works representing these various positions, and intermediate ones, are as follows: Daniel Dennett, 'Quining Qualia', in A. Marcel and E. Bisiach (eds.), *Consciousness in Contemporary Science*, (Oxford: Oxford University Press, 1988); idem, *Consciousness Explained* (Boston: Little, Brown, 1991); Fred Dretske, 'Conscious Experience', *Mind*, 102 (1993), 263–283; idem, *Naturalizing the Mind*; David Rosenthal, 'Two Concepts of Consciousness', *Philosophical Studies*, 49 (1986), 329–359; John Searle, *The Rediscovery of Mind* (Cambridge, Mass.: MIT Press, 1992); Joseph Levine, 'On Leaving Out What It's Like', in M. Davies and G. Humphreys (eds.), *Consciousness*, (Oxford: Blackwell, 1993); Martin Davies, 'Externalism and Experience', in A. Clark, J. Ezquerro, and J. M. Larrazabal (eds.), *Philosophy and Cognitive Science: Categories, Consciousness, and Reasoning*, (Dordrecht: Kluwer, 1996); Charles Siewart, *The Significance of Consciousness* (Princeton: Princeton University Press, 1998); David Chalmers, *The Conscious Mind: In Search of a Fundamental Theory* (Oxford: Oxford University Press, 1996). For further discussion that hinges

The variety of theories about consciousness, and the depth of disagreement among theorists, suggests that there may be differences at the intuitive level—differences about the explanandum. Block suggested fruitfully that there are at least two notions of consciousness. One has to do with phenomenal, felt, qualitative aspects of experience. Another has to do with accessibility to rational deliberation, or at least to reflection and use in verbal and other rational enterprises.[46] Whether or not this view is correct, it seems to me to have engendered greater awareness of the complexity and elusiveness of the phenomena that have been discussed in philosophizing about consciousness.

Discussions of consciousness have opened up what had been a neglected, almost taboo subject in philosophy. This has been a good thing. On the other hand, it seems to me that progress toward genuine understanding has been at best mixed. The difficulty of the subject has provoked not only a number of wildly implausible theories. It has also encouraged what seems to me to be some backsliding in methodology and clarity of discussion in many presentations in this area. Metaphors, appeals to disputed and sketchily described introspection, unexplained terms, hastily and poorly explained technicalia, and grandiose programs have been more prominent than is good for a subject. These are early days. A certain amount of floundering is inevitable in initial work on a difficult topic. It may well be that a deeper scientific grip on relevant areas of neural, sensory, and cognitive psychology will be necessary before impressive progress emerges. It is well also to remember that many of the most fundamental aspects of our experience of the world—aspects that throw up some of the most basic and long-standing philosophical problems—have a way of remaining puzzling, despite progress in associated sciences. Consciousness may remain a case in point.

I want to close by summarizing some of the main changes in *both* philosophy of mind and philosophy of language in the second half of the twentieth century. Three major, possibly durable contributions in these areas during the period are the criticism of the positivist theory of meaning; the development of a vastly more sophisticated sense of logical form, as applied to natural language; and the fashioning of the nondescriptivist account of reference, with the extension of the line of thought associated with this account into the philosophy of mind. Different philosophers would, of course, provide different lists of achievements, given their own sense of what is true and important.

The dominant currents during the period are more easily agreed upon. The central event is the downfall of positivism and the reopening for discussion of

on more technical and methodological issues, see Ned Block and Robert Stalnaker, 'Conceptual Analysis, Dualism, and the Explanatory Gap', *The Philosophical Review*, 108 (1999), 1–46; David Chalmers and Frank Jackson, 'Conceptual Analysis and Reductive Explanation', *The Philosophical Review*, 110 (2001), 315–361. A good anthology is Block *et al.* (eds.), *The Nature of Consciousness*.

[46] Ned Block, 'On a Confusion about a Function of Consciousness', *The Behavioral and Brain Sciences*, 18 (1995), 227–247. Cf. Tyler Burge, 'Two Kinds of Consciousness' in Block *et al.* (eds.), *Nature of Consciousness* (Ch. 17 above).

virtually all the traditional problems in philosophy. This event was accompanied by the rediscovery of Frege, the application of logical theory to language, and the rise of the philosophy of language both as a preliminary to reflection on other subjects, and as a more nearly autonomous discipline. The computer paradigm and complex outgrowths of the philosophy of language have brought the philosophy of mind to dominance in the last decade.

Positivism left behind a strong orientation toward the methods of science. This orientation has fueled the acceptance of materialism in the philosophy of mind and, somewhat belatedly, the development of areas of philosophy (philosophy of physics, mathematics, biology, psychology, linguistics, social science) that take the specifics of scientific theories and practices into account.

For all this, the main direction of philosophy during the period has been toward a broader-based, more eclectic, less ideological approach to philosophical problems—and a greater receptivity to interplay between modern philosophy and the history of philosophy. Philosophy of mind emerged as an area of intense ferment not simply as a product of interaction between philosophy and such disciplines as psychology and linguistics. That ferment also represents a greater interest in traditional questions, questions about what is morally and intellectually distinctive about being human. It is hard to overemphasize the degree to which leading North American philosophers have since the 1950s broadened their sympathies toward traditional questions that still help frame what it is to lead a reflective life.

This broadening seems not to have seriously undermined the standards of rigor, clarity, and openness to communal check bequeathed by such figures as Frege, Russell, Carnap, Hempel, Gödel, Church, and Quine. Partly because of its close connection with the development of mathematical logic in this century, the standards of argument in philosophy have certainly been raised.

A corollary of this change, and of the personal example of the positivists in carrying on open, dispassionate discussion, has been the emergence of philosophical community. One of the glories of English-speaking philosophy in the last half-century has been the fruitful participation of many philosophers in the same discussions. Unlike much traditional philosophy and much philosophy in other parts of the world, English-speaking philosophy has been an open, public forum. The journals of the field, including notably *The Philosophical Review*, bear witness to a sharing of philosophical concerns, vocabularies, and methods of dispute. We now take this sharing for granted. But in historical perspective, it is remarkable. Although I think that philosophy is not and never will be a science, it has taken on this much of the spirit of science. That is, to my mind, the more important achievement.

This overview has provided at best a blurred glimpse of the enormous complexity and variety of discussion in philosophy mind during the last half-century. It is deficient as a picture not only in its oversimplifications and limited scope, but also in its failure to convey the life and nature of the animal. Philosophy is not primarily a body of doctrine, a series of conclusions or systems or

movements. Philosophy, both as product and as activity, lies in the detailed posing of questions, the clarification of meaning, the development and criticism of argument, the working out of ideas and points of view. It resides in the angles, nuances, styles, struggles, and revisions of individual authors. In an overview of this sort, almost all the real philosophy must be omitted. For those not initiated into these issues, the foregoing is an invitation. For those who are initiated, it is a reminder—a reminder of the grandeur, richness, and intellectual substance of our subject.