24. ANTI-REDUCTIONISM AND THE MIND-BODY PROBLEM

CLAUDIA M. MURPHY
UNIVERSITY OF COLORADO, COLORADO SPRINGS

ABSTRACT. I argue that there are good reasons to deny both type-type and token-token mind-brain identity theories. Yet on the other hand there are compelling reasons for thinking that there is a causal basis for the mind. I argue that a path out of this impasse involves not only showing that criteria of individuation do not determine identity, but also that there are sound methodological reasons for thinking that the cause of intelligent behavior is a real natural kind. Finally, a commitment to this methodology suggests both that these familiar anti-reductionist arguments fail to establish that identity is impossible and at the same time suggest that the preferred alternative will be some version of neutral monism.

Anti-reductionist arguments typically appeal to the principle of the indiscernibility of identicals. With reference to mind-brain identity theories, the argument might involve the claim that properties of the mental are not properties of the physical and vice versa. A more contemporary statement of this argument is to the effect that type-type identifications are impossible because criteria which sort the mental into kinds are very different from criteria which sort the physical into kinds.1 And one can extend the argument in suggesting that even token-token identity seems implausible because criteria which individuate the mental are very different from criteria which individuate the physical.

So, it seems as if there are good reasons to deny both type-type and token-token identity between the mental and the physical. Yet, on the other hand there are compelling reasons for thinking that there is a causal basis for the mind. In what follows, first, I will explore the reasons for rejecting both type-type and token-token identity. Second, I will explore some reasons for thinking that there is a causal basis for the mind. Third, I will argue that a path out of this impasse will involve not only showing that criteria of individuation do not determine identity; but also that there are sound methodological reasons for thinking that the cause of intelligent behavior is a real natural kind. And finally, a commitment to this methodology suggests both that these familiar anti-reductionist arguments fail to establish that identity is imposs-
ble and that the preferred alternative will be some version of neutral monism.

**PART I: REASONS TO DENY TYPE-TYPE IDENTITY**

In order to understand an identity theory, we first need to look at the supposed differences between mental and physical phenomena. Mental phenomena fall into two types: sensations and propositional attitudes. Initial objections to materialism focused on sensations as the major obstacle to a successful reduction. Lately, however, the opposition has centered on the realm of the intentional, with the understanding that reduction is even more problematic for propositional attitudes than for sensations. Hence I shall confine my remarks to propositional attitudes.

A distinction which dates back to the Middle Ages is that mental phenomena enjoy intentional inexistence whereas physical phenomena do not. Brentano more fully developed the idea by claiming that the mind has a unique ability to be directed toward objects that need not exist. This is the property of intentionality. Straightforward physical events do not seem to have this property since for two physical events to be related to one another both must exist.

Type-type identity theories have been developed in order to explain the connection between the mind and the body. They rely on the following assumption: It is possible to explain why a person is in a particular mental state (or engaging in meaningful behavior) by reference to some physiological mechanism identical with that mental state. I plan to examine two arguments for the conclusion that this assumption is unintelligible. First, in *ESP and Psychokinesis* Stephen Braude argues that this assumption can be undermined by showing that it rests on one or two false presuppositions:

1. The theorist Platonically assumes that a mental state can be classified by finding necessary and sufficient conditions for being a mental state of that type.

2. The theorist supposes that a physiological process can be functionally unambiguous, i.e., that structurally identical brain states will have identical function or representational properties.

Let's examine the first assumption. In the natural realm it seems as if kinds are individuated by means of necessary and sufficient conditions. Water, for example is categorized by reference to chemical composition. But the assumption that mental states are also so classified has led to numerous philosophical problems. Braude, following Wittgenstein, argues that mental states are classified positionally rather than by means of necessary and sufficient conditions.

Recall that a thought of a certain kind (say the thought of an old rabbi) may take an indefinite number of forms, having nothing more of relevance in common than the fact that they are of that kind. Thus when I am thinking of an old rabbi, I may be remembering a verbal description, or having
an image of an old rabbi, I may be remembering a verbal description, or having an image of a person, or perhaps a more complicated image (say, of a rabbi conducting a bar mitzvah), or perhaps I am just presenting to myself the words "old rabbi", or perhaps I am imagining these words sung by the Mt. Sinai Hospital Chorale. And of course each of these categories will have innumerable variations as well. The verbal descriptions or the images (for example) can differ subtly or dramatically from one another.\(^3\)

Mental states are not classified by means of necessary and sufficient conditions: they can differ in a indefinite number of ways. So the first presupposition looks false. Braude argues that the second presupposition is also false. Take, for example, the crudest conception of a mental event—a mental image. According to the neurophysiologist, the mental image is identical with a type of neurophysiological structure or process in the brain. Braude argues that this presupposes that a physiological process can be functionally unambiguous. To convince us that this presupposition is false, Braude gives a concrete example which is analogous to the mental and the physical. Consider written language. The statements in language are analogous to the mental in that they too have the property of intentionality. And written languages allow us to draw a fairly straight-forward distinction between syntax and semantics. Syntactically, statements are strings of marks, e.g., inkmarks, chalkmarks. Semantically, statements represent or mean. Braude argues that we do not classify syntactic units in the same way that we classify semantic units. Consider the statement "The bill is large." Although it is only one syntactic unit, its meaning is ambiguous. The word "bill" could refer to a monetary unit, the beak of a bird, or to the statement of a debt. On the basis of this example, one can conclude that syntactically identical strings of marks need not be identical semantically. Grasping the other side of the horn, one can give examples in which syntactically different strings of marks are identical semantically. Consider "It's raining" and "Es regnet." Braude would cash in the analogy as follows: identical brains states need not be identical semantically. So type-type identity seems doomed, given that physiological processes are functionally ambiguous.

Some functionalists, such as Block and Fodor, have advanced an argument for a related conclusion. They argue that type-type identification seems very implausible. This follows from the fact that the mind is to be understood in terms of an abstract organization of functional states. Since a variety of diverse physical mechanisms could instantiate this organization, it seems mistaken to reduce a functional account to principles embodied in any one physical instantiation. They argue that this shows that type-type identity theories are probably false, and that the burden of proof is on the materialist.

If someone suggests a putative physical state common to all and only realizations of S, it is a simple matter to dream up a nomologically possible machine that satisfies the machine table but does not have the designated physical state. Of course, it is one thing to say this and another to prove it, but the claim has such overwhelming prima facie plausibility that the burden of proof is on the critic to come up with reasons for thinking otherwise.\(^4\)
The point seems to be that the criteria for identifying an event via its functional role are significantly different from criteria for identifying an event in terms of its physical structure.

PART II: REASONS FOR DENYING TOKEN-TOKEN IDENTITY

Donald Davidson, in "Mental Events," introduces another identity theory which may be immune to the above criticisms. Davidson argues that types of mechanism are not identical with types of mental state for reasons similar to Braude's. They agree that psychology and physiology do not carve the world up into kinds in the same way.

There is no assigning beliefs to a person one by one on the basis of his verbal behavior, his choices, or other local signs no matter how plain and evident, for we make sense of particular beliefs only as they cohere with other beliefs, with preferences, with intentions, hopes, fears, expectations, and the rest. It is not merely, as with the measurement of length, that each case tests a theory and depends upon it, but that the content of a propositional attitude derives from its place in the pattern.

Davidson continues by defending another theory which at first seems immune to this criticism. He argues that even though there may not be identity between types, nevertheless token of mental events may be identical with tokens of physical events. Davidson calls this view "Anomalous Monism." The monism is anomalous because there will not be any psychophysical laws correlating types of mechanism with types of mental state.

But Braude rejects even this weaker kind of identity theory. He points out that this view will have absolutely no explanatory power. If it does not matter what kind of brain state is identical with a particular thought, Braude claims, then it could be any neurophysiological structure. For this reason, Braude contends that anomalous structure. For this reason, Braude contends that anomalous monism is actually contrary to the spirit of materialism. The hope was that an explanation in physiological terms would explain why a person was having one thought rather than another. Braude concludes: "But these explanatory and predictive features of a materialist theory may be secured only if there are type-type correlations between the mental and the physical."

Block, from a functionalist point of view, also suggests that token physicalism is a much weaker thesis than the one to which materialists have been committed. He further points out that some functionalists admit that a nonphysical state could play the same role as a mental state.

... as functional theorists have often pointed out, a nonphysical state could conceivably have a causal role typical of a mental state. In functional specification terms, there might be a creature in which pain is a functionally specified soul state. So functionalism opens up the possibility that even if our pains are physical, other pains might not be. In light of this point, it seems that the support that functionalism gives to even token physicalism is equivocal.
So, according to Block, it is not even clear that functionalism can endorse any variety of materialism.

Bruce Goldenberg makes an even deeper criticism in "A Problem with Anomalous Monism." He argues that Davidson assumes without reason that referring expressions in a sophisticated neurophysiology will have the same extensions as the expressions we currently use to pick out mental events. His point is that since Davidson admits that psychology and physiology do not divide the world into the same types, Davidson needs some further reason for assuming that these radically different criteria of individuation will yield the same individuals.

So the conclusion is that both type-type and token-token mind-brain identity theories are false or at least uninteresting. Braude suggests as an alternative that we abandon the search for an explanatory unit more fundamental than the person. He proclaims that nature has no intrinsic structure and so no parsing of objects is to be preferred over any others. So whenever a mode of explanation rests on false or absurd presuppositions, Braude thinks that certain facts should be recognized as ultimate or as needing no explanation, and that that mode of inquiry should be abandoned. Braude adds that if he is right, then cognitive psychologists, generative linguists, neurophysiologists, etc. should all be unemployed.

PART III: THERE ARE REASONS TO BELIEVE THAT THERE IS CAUSAL BASIS FOR MIND

An assumption which undergirds many recent explanations of human cognitive abilities is that a physical mechanism can explain intentionality. Cognitive psychologists, neurophysiologists, generative linguists, and as I mentioned, many philosophers assume an explanatory physical mechanism. This assumption can take either of two forms: the mind is identical with the brain or the mind and the brain interact in a way which can be given a significant causal explanation. Thus not only do identity theorists seem committed to this assumption but also any dualist who desires a causal explanation of the interaction between the mind and the body is similarly committed. Even Descartes insisted that minds are embodied, so the issue of causal interaction for a dualist becomes acute whenever the world impinges on humans or humans impinge on the world.

But not only philosophers seem to depend on such an assumption. Memory trace theorists, for example, assume that it is possible to remember an event only due to a trace within the person—maybe a modification of the brain—caused by an association with that event. The various versions of the theory are an attempt to provide a scientific answer to questions such as "When a person whistles the 'Fifth Brandenburg Concerto', how can he or she possibly get it right?" Roughly, such questions are to be answered as follows: the central nervous system is analogous to an extremely powerful, very sophisticated computer. When one whistles a tune he or she remembers, some biological mechanism which works like a high fidelity recorder has produced a trace or engram of the tune in question.
Consider another example. Some brain scientists think that it is theoretically possible to predict that a person is having a particular thought by knowing the arrangement of neurons in the brain. This assumption, although it can take various forms, seems to undergird neurophysiology; it is not questioned.

So it seems as if there are both empirical reasons and prima facie philosophical reasons for desiring a causal explanation for the basis of the mind. I think that one could further argue that there are important psychiatric and medical reasons for endorsing this type of research program. In Part IV I will consider methodological reasons for espousing neutral monism.

PART IV: THE PATH OUT OF THE IMPASSE

I will argue that the above mentioned criticisms of an identity theory fall short of the conclusion. My major point will be that even if there were to be some important differences in criteria of individuation between the mental and the physical, nevertheless these criteria do not dictate the identity of an object. So even though it may be true that a dualistic psychology and physiology do not divide the world into kinds or even individuals in the same way, it need not follow that any assumption of identity is false.

Let me begin by looking at an example where even token-token identity seems impossible. I will consider an example which involves mutually incompatible concepts, i.e., concepts which cannot be true of the same object in the same respect.

The mind-body problem is an instance of a larger philosophical issue concerning the relations between ordinary visible objects and microphysical objects. So examining seemingly unrelated examples of conceptual controversy might illuminate the issue. Looking at the history of science reveals that rival theorists have sometimes required conceiving of the "same thing" in mutually incompatible ways. For example, the particle theory of light requires us to conceive of light as consisting of tiny particles of matter. Whereas the wave theory of light requires us to conceive of light as electromagnetic waves. These waves are thought to be perpendicular fields undulating through space; but a field is determined by plotting the force a particle would experience if it were there. Hence electromagnetic waves are really hypothetical motions rather than material substances. And nothing can be both a substance and a hypothetical motion in the same respect. One might say that the particle and wave theories do not divide the world into individuals in the same way.

T.S. Kuhn and P.K. Feyerabend have used examples like this to argue that in such cases rival theorists are not even talking about the same things. They suggest that important classes of concepts do not preserve meaning (or reference) across theory change. And they argue that since there is no neutral vocabulary with which to discuss the merits of rival theories, theory change in science is no more rational than religious conversion.
If these anti-reductionist arguments against the identity theory are correct, then it seems that concepts of the mental and concepts of the physical are also mutually incompatible. The two theories I have in mind include neurophysiology and folk psychology (as the Churchlands use this term.) The claim is that physiology and folk psychology are incommensurable theories—the theorists are not even talking about the same things; identity seems impossible.

In response to this type of difficulty, Hilary Putnam introduced a causal theory of reference for terms denoting natural kinds. By re-examining what it is for a concept to refer to an object, Putnam argues that mutually incompatible concepts may refer (de re) to the same thing although opposing theorists need not share any concept of the disputed subject.

Putnam begins by challenging two assumptions that have traditionally been made in the theory of meaning.

1. Knowing the meaning of a term is just a matter of being in a certain psychological state.

2. The meaning of a term determines its extension. Same-ness of intension (sense) entails sameness of extension (referent).

Putnam points out that, since Frege, philosophers have agreed that two concepts could have different senses although they shared a referent, but the opposite was thought to be impossible. In other words, it seems impossible for two concepts to share a sense although they have different referents. Putnam challenges these assumptions by arguing that they are not jointly satisfied by any notion of meaning. Through the use of examples he tries to persuade us that two speakers can be in the exact same psychological state even though the extension of a concept of one of the speakers differs from the extension of the same concept of the other speaker. Since by hypothesis they share the same psychological state, the extension of a concept is not determined by the psychological state of the speaker.

Putnam's major example is a science fiction story about a planet called twin earth, which is exactly the same in all respects to Earth except the chemical composition of what they call "water," which is XYZ instead of H2O. Putnam proposes that we consider a typical English speaker in about 1750 and his doppelganger on Twin Earth. He claims that the psychological states of the "twins" are identical, although the extensions of the concepts differ. "Water" on earth is not a concept of the same thing as "water" on twin earth because the extension of "water" on earth is H2O, whereas the extension of "water" on twin earth is XYZ. So his conclusion is that the psychological state of the speaker does not determine the extension of the concept.

In claiming that the psychological state of the speaker doesn't determine the extension of a concept, Putnam opens the door for an alternative theory of meaning. His main project is to re-explain how the referent of a term is fixed. Following Kripke, he argues that the referent of a concept is fixed by a causal chain that links uses of that concept with certain original or paradigmatic instances. Then he claims that the referent of a concept is all that that which has the same es-
sence as the original instance. What the essence is, is to be determined by scientific inquiry. So the essence is totally independent of the meaning of the concept. The referents of a concept need not even be things of which the concept is true.

This view seems attractive because we can think of the history of science as the history of attempts to refer to natural kinds and reveal their essences. Putnam's response to Kuhn and Feyerabend is that they were operating with the wrong notion of meaning. Although it is true that the particle and wave conceptions of light do not have the same sense and are not true of the same kinds of things, nevertheless they can both be conceptions of the same thing since both concepts were introduced in connection with instances which have the same essence.

I think that the Putnam-Kripke theory of meaning is inadequate for a variety of reasons. However, one fairly easy way to suggest this is to consider concepts which are not well explained by the view. In particular, Putnam's view turns out not to work well for concepts which do not refer to anything, e.g., demons, phlogiston, etc. Since there is nothing to which these concepts could be said to refer, there is nothing of which these concepts are misconceptions. So according to Putnam, none of these concepts are concepts of anything. And in fact they all have exactly the same meaning.

Roughly, I have been introducing T.L. Short's solution to the problem in "An Analysis of Conceptual Change." Short argues that theorists with mutually incompatible concepts can be referring to the same thing if each has a concept which is intended to be true of the presumed extension of some third, prior concept. He thinks that concepts which play the role of prior concepts in these scientific controversies are generally those in which the phenomena are thought of indirectly, by their presumed relation to something else. This vagueness or indirectness allows speakers to identify an object in a manner which is compatible with subsequently developed mutually incompatible concepts.

Short thinks that this manner of concept introduction is inspired by what Peirce called hypostatic abstraction. One example of this type of concept introduction would be a speaker postulating something real which is supposed to account for familiar phenomena. She is doing so in an indeterminate or abstract manner, identifying this real thing as that which causes or is responsible for the superficial phenomena. Short points out that this process of abstraction is of a special kind. Usually when one goes from a specific concept to the more abstract, one gets only a more general concept and not a less specific concept of the same type. An example Short gives is as follows:

Thus, if we omit the specification "unmarried" from the explication of the concept of a bachelor, what we have is the explication of the concept of an adult human male. But we have seen in Section 5, that the intensional object of a hypostatic abstraction can be quite a narrow type even though the characteristics of that type are not specified in the abstraction.

Short's main insight is that pre-scientific concepts are hypostatic abstractions which are compatible with any scientific theorizing. After the phenomena have been identified via a hypostatic abstraction, scien-
tists can disagree about the specific nature of that presumed natural kind. So, contra Kuhn, rational discussion of competing theories seems possible. Short summarizes his response as follows:

On the view presented here, alternative theories about the same things, basic or not, remain incommensurable. They allow us to view these things in quite different ways, requiring something like a gestalt switch as we pass from one to another. But opposing theories are not incommensurable in the sense that there are no concepts, compatible with both theories, through which to identify their common subject matter.\(^{16}\)

For example, Short might claim that the particle conception of light is incommensurable with the wave conception of light. But the prescientific conception—the hypostatic abstraction—is compatible and hence commensurable with both rival, more sophisticated conceptions.

Initially, Short's view seems close to Putnam's. Putnam would claim that since the particle and wave conceptions of light are linked by causal chains to an original or paradigmatic instance, both are conceptions of the same thing. But Putnam also argues that the causal chains which fix the referent of the concept are not represented in the concept itself. Hypostatic abstractions, on the other hand, contain within their explications descriptions, however vague, of the causes of the familiar phenomena.

A way to clarify this disagreement involves arguing that Short's position seems tenable, whereas Putnam's does not. Putnam, along with Kripke, seems to think that the reference of the original concept gets fixed by a baptismal ceremony. In "The Meaning of Meaning," Putnam appeals to the device of an ostensive definition. The original baptiser points at something and gives it a name. The difficulty with the view, of course, is that ostensive definitions are not unambiguous. Pointing in itself would not pick out a natural kind as opposed to one of its properties. And pointing would not single out the essence of the object. In fact an object does not instantiate a single natural kind. So the original baptism would not help us to decide when another object should or should not be called by that same name, unless we know what the baptiser meant.

Short, on the other hand, does not rely on ostensive definition. A hypostatic abstraction, although vague, is a concept of that which causes the familiar phenomena. So if there is a real natural kind which causes the familiar phenomena, then the pre-scientific theorizer has succeeded in referring to that natural kind.\(^{17}\)

**PART V: NEUTRAL MONISM**

Short's insight is helpful only in those case in which rival theorists in fact are referring to a real natural kind through a hypostatic abstraction. Could this have been the case in relation to conceptions of the mental and the physical? If so, then Short's insight can be applied to the mind-body problem. In order to motivate this suggestion, I plan
to construct a pre-scientific conception of the cause of intelligent behavior. Needless to say, this conception will be speculative.

In "Empiricism and the Philosophy of Mind," Wilfrid Sellars indulges in what he calls anthropological science fiction in order to explain how humans learn to speak of inner episodes and immediate experiences. Sellars asks us to image a Rylean language which allows these humans to speak of public objects in space and time. The logical apparatus at hand permits them conjunction, disjunction, negation, quantification and the conditional. And finally this language includes the resources of semantical discourse, the concepts of meaning, truth, reference, etc. But note that these Rylean ancestors have no conception of inner episodes or immediate experiences. They think although they have no concept of thought.

An enrichment of the Rylean language is theoretical discourse. Sellars sees these individuals as postulating abstract, vague causes for intelligent behavior.

It will not surprise my readers to learn that the second stage in the enrichment of their Rylean language is the addition of theoretical discourse. Thus we may suppose these language-using animals to elaborate, without methodological sophistication, crude, sketchy, and vague theories to explain why things which are similar in their observable properties differ in their causal properties, and things which are similar in their causal properties, differ in their observable properties. Sellars' description (up to this point) sounds remarkably like Short's conception of a hypostatic abstraction. Consider the following behaviors. At t₁, Jones says 'I calculate' and then writes '2+2=4'. At t₂ Jones simply writes '2+2=4'. Our Rylean observers note that the two inscriptions although similar, differ in their causal antecedents. The first inscription seems to be caused by the verbal utterance "I calculate." So in this case, they postulate a real cause which accounts for the second inscription. It is important that the nature of this real cause not be specified by their conception; they are simply identifying a real natural kind which is causally responsible for intelligent behavior.

Some years later we can imagine that the individuals begin developing theories about the specific nature of that presumed natural kind. One theorist may have hypothesized that these causes are mental events, they lack spatial location, they have phenomenal properties such as nagginess, they are incorrigible, etc. Another theorist, in the opposite camp, has hypothesized that these causes are nothing more than blood flowing through contiguous cells in the brain. At this point, theorists have developed mutually incompatible concepts. Eventually these two views evolve into Folk Psychology and neurophysiology. The theories although incommensurable with each other are each commensurable with the prior, vague conception of the cause of intelligent behavior. But even though Folk Psychology and neurophysiology are incommensurable with each other, it doesn't follow that any hypothesis of identity is either false or unintelligible. If the cause of intelligent behavior is a real natural kind, then some form of identity may be true.
The anti-reductionist arguments I have examined only apply to our current ways of carving up the world. However, nothing in these arguments is incompatible with our developing a more adequate vocabulary (perhaps neither strictly mental nor strictly physical) for understanding the nature of intelligent behavior. What philosophical reasons are there for thinking that there is a causal basis for the mind that is a real natural kind? First, returning to the analogy with conceptual change in science, I think that there are some examples that illustrate the sort of change that I am looking for. At times in the history of science rival theorists have employed mutually incompatible concepts to discuss what had been identified as one kind of thing by means of a hypostatic abstraction. It is certainly possible for them to discover that they really were not talking about one natural kind after all. In other words, one kind of conceptual change involves further discrimination between natural kinds. But it also seems possible that theorists with incompatible concepts are nevertheless talking about only one natural kind. This sort of conceptual change seems to be in process in the particle-wave controversy. Physicists seem committed to claim that light is a real natural kind that sometimes acts like a particle and sometimes acts like a wave. (I'm not convinced of the coherence of the current conception, but perhaps further investigation will yield a conception that is neither particulate nor undulatory.) There are also examples of this type of conceptual change in the history of medicine. At times illnesses which seemed quite different turned out to be the same disease.

In any case it seems to me that there are no empirical reasons for preferring one of these types of conceptual change rather than the other. Are there any methodological reasons? I've hinted at an answer to this question in my earlier discussion of Sellars. He argues that conceiving of the cause of intelligent behavior as a theoretical natural kind is a way out of the impasse between dualism and behaviorism as regards the acquisition of concepts of the mental. He thinks that the dualist is correct that humans have privileged access to their thoughts and sensations, yet the behaviorist is correct that our language is a public one. He argues that one can only reconcile these two insights if she is prepared to say that the concepts of the mental were originally introduced as concepts of theoretical entities as causes of intelligent behavior. The reporting role of such concepts builds on and presupposes their public role. So it seems as if there are sound methodological reasons for preferring neutral monism to dualism or materialism.

I also think that a neutral monism is to be preferred over any version of eliminative materialism. Several philosophers, (Rorty, Feyerabend, Paul Churchland, etc.) have argued that our common sense conception of the mental is simply false, and will be replaced by an adequate neurophysiology. Rorty seems to be defending this view only as regard our common sense conceptions of sensations, but Churchland extends his view to include our conception of all intentional phenomena. Churchland claims:

But FP (Folk Psychology) is no part of this growing synthesis. Its intentional categories stand magnificently along, without visible prospect of reduction to that larger corpus.
A successful reduction cannot be ruled out, in my view, but FP's explanatory impotence and long stagnation inspire little faith that its categories will find themselves neatly reflected in the framework of neuroscience. On the contrary, one is reminded how alchemy must have looked as classical mechanics was being articulated, or how the vitalist conception of life must have looked as organic chemistry marched forward. . . . what we must say is that FP suffers explanatory failures on an epic scale, that it has been stagnant for at least twenty-five centuries, and that its categories appear (so far) to be incommensurable with or orthogonal to the categories of the background physical science whose long-term claim to explain human behavior seems undeniable. Any theory which meets this description must be allowed a serious candidate for outright elimination.

This strategy seems similar to throwing out the baby with the bath water. If we should simply eliminate our common sense conceptions of the mental, then it is extremely difficult to understand what counts as a successful explanation of the mind. Where does one begin? What constraints can be placed on rival theories?

Short has provided another alternative. If the cause of intelligent behavior is a real natural kind, then while concepts of the mental can be incommensurable with concepts of the physical, nevertheless both can be commensurable with a much less specific, prior concept. And if there are good reasons for thinking that the cause of intelligent behavior is a real natural kind, then it seems as if neutral monism is to be preferred.

ENDNOTES

1 e.g., Donald Davidson, "Mental Events," in L. Foster and J.W. Swanson (eds.) Experience and Theory Amherst: University of Massachusetts Press 1970.


3 Ibid., 190.


5 Davidson, op. cit., 96.

6 Braude, op. cit., 183.

7 Block, op. cit., 181.

This example is developed to a much greater length (and criticized for the presuppositions involved) by H.A. Bursen, in *Dismantling the Memory Machine*, Dordrecht: D. Reidel, 1978.


Putnam has since changed his views about reference.


Ibid., 308.

I have some serious misgivings about the adequacy of Short's view as a general account of conceptual change in science. A serious problem that I see with his view is that he is committed to a picture of science as a history of references to natural kinds and attempts to reveal their essences. This view seems to preclude the introduction of any genuinely new theoretical frameworks.


Ibid., 183.


Churchland, *op. cit.*, 75 & 76.