
PHYSICALIST THEORIES OF MIND

Behaviorism

Given the troubles of dualism, one may be tempted by what is easily the most straightforward solution to the mind-body problem: physicalism. On this view, everything that exists is physical; so the mind is a physical thing, if it is a thing. If proposition (1), that the mind is a nonphysical thing, is false, which it is if everything is physical, then the mind-body problem is solved. The mind is a physical thing, and so there is nothing to stop it from interacting with other physical things, including the body. It remains true, however, that physical and nonphysical things, on this view, cannot interact. But it doesn't matter, since there are no nonphysical things.

Well and good, but *in what way* is the mind supposed to be a physical thing? There are a number of different possibilities.

Behaviorism is the view that the mental is the behavioral. Mind is behavior. The mind is the body, considered from the point of view of its behavior. Some hardline behaviorists actually went so far as to deny the existence of the mind and mental events, over and above behavior. There is no mind, but only behavior. This is a very simple but pretty extreme point of view that has not found much favor among philosophers or scientists recently. Part of the problem is that we do seem to be acquainted with our own mental states, our thoughts and feelings, and they are not nothing at all. Another part of the problem is that there do seem to be obvious examples of an interaction from mind to body.

A second and more reasonable version of behaviorism took the line that, from a scientific point of view, we should not study the mind and mental events, because they cannot be directly observed; their existence must be inferred from the external behavior of human subjects. This is not the strongest line of thought, it has to be said, since many entities studied in science cannot be observed directly, but we infer their existence from their effects. Electricity is an example. We know about it by watching lightning, for example, or by understanding Maxwell's equations, or how a radio works.

Nevertheless, one can understand how, in the atmosphere of the religiously oriented dualism that prevailed in philosophy at the beginning of the twentieth century,

and which many scientifically oriented people found uncongenial, the bold claim could be advanced, on behalf of psychology, that science should allow as its subject matter only what can be directly observed. This is certainly very different from saying that its subject matter does not exist.

An even more reasonable variant of behaviorism is that mind as such is not interesting or important, and its study should be replaced by the study of behavior. There is no mention in this view of what is directly observable. It is almost like saying, "I am more interested in behavior than I am in mind." This is, of course, an impossible view to rebut, if it is true that you are more interested in behavior than in mind; but the question remains whether you *should* be more interested in mind as such, or whether its study would offer you some benefit.

This third and more reasonable line of thought, however, would make it impossible to solve the mind-body problem in a way that is satisfactory for science, or even to *state* it. We should not study or talk about minds, so we will never be in a position to say either that the mind is a nonphysical thing with any scientific authority, or, for that matter, that it is a physical thing. Our first proposition, that the mind is nonphysical, is one whose truth or falsity we should not actively pursue, because its truth or falsity is something that should not be talked about!

There is a fourth form of behaviorism, however, that is more appealing than any of the first three forms. It is the view that the mind is behavior in the sense that any proposition about the mind can be “translated” into a proposition about behavior. So, for example, if I say “I am tired,” I am reporting not the presence of an inner feeling of drowsiness, but rather of a tendency or disposition to stop work, to lie down, to close my eyes perhaps, to rest, and so on. All of these things are external behavior, observable by others and fully within the purview of science and of common observation.

Gilbert Ryle wrote in his influential 1949 book *The Concept of Mind* that

when we describe people as exercising qualities of mind, we are not referring to occult episodes of which their overt acts and utterances are effects; we are referring to those overt acts and utterances themselves.¹

It is hard to believe, reading the admittedly rather few passages like this in his book, that Ryle was not a behaviorist, and indeed he himself remarked of the book that when he wrote it, “certainly one of my feet was pretty firmly encased in this boot.”² Nevertheless, there is more to the story. Ryle writes in the passage above that when we talk about minds, “we are not *referring* to occult episodes” (my

emphasis); but there is a case to be made that all the same he does not deny the *existence* of these episodes. Perhaps he means that when we say publicly that a person is tired, we are “referring,” not to that person’s private and inner feeling of tiredness hidden from others, but rather to his tendency or disposition to stop work, to lie down, to close his eyes, to rest, and so on. This is not to deny that the inner feeling exists. In chapter 6 I describe the other side of Ryle’s view, his “dissolutionism” as it has been called, and again take up the question whether he is to be considered a full-blooded behaviorist.

What is wrong with the idea that the mind just *is* some sort of behavior? One difficulty is that this view seems to leave out what we think of as the “inner” life of thoughts and feelings—the mind! Behaviorism solves the mind-body problem by denying the mind in one way or another. We can produce behavior without it, and without its rich experience of sensation and perception, colors, sounds, and tastes, for example, or qualia. We can easily imagine a machine that reacts to red things just as we do, picking them and eating them, perhaps, but which has no experience of the colors. It behaves as if it saw red, but it does not have the experience. This has been called the “problem of absent qualia.”

There are other twentieth-century philosophers who, like Gilbert Ryle, might also seem to be offering behaviorist arguments but are not. A famous example is the

“beetle-in-the-box” part of the so-called private-language argument in Wittgenstein’s *Philosophical Investigations*. Wittgenstein is arguing that there could be no language in which we could report our own private sensations. Suppose, he writes, that everyone has a box with something in it, or perhaps nothing at all. There is a rule that no one is allowed to look inside anyone else’s box. Everyone calls what is in his own box a “beetle.” But where no checking is allowed about what is in anyone else’s box, the word “beetle” would not come to mean “an organism of the order *Coleoptera*, with hard fore-wings,” but rather “whatever is in anyone’s box.” Yet Wittgenstein explicitly denies that he is trying to deny the existence of sensations, somethings in the boxes. The issue is one of meaning.

Behaviorism does indeed solve the mind–body problem, very easily, by denying that the mind is a nonphysical thing. Behaviorism simply denies proposition (1). So the discussion at this point should turn to the question of how plausible behaviorism itself is. The judgment of history, it is fair to say, is “Not very.” One powerful reason is the problem of absent qualia, mentioned above. Another objection is the possibility of the *inverted spectrum* and its analogues in other sensory modalities. We can imagine people behaving systematically in the right way, but having the “wrong” experiences. Their “inner experience” might be of all the colors, but with their positions in the visual field reversed

Behaviorism does
indeed solve the
mind–body problem,
very easily, by denying
that the mind is a
nonphysical thing.

from ours. The subjects with inverted color experiences would see a cyan green-blue color where we see red, a blue where we see orange, and so on throughout color space. But the *behavior* of these people would be the same as ours. When we see red, and call it “red,” they see what we call “cyan,” and call it “red,” and when we see cyan, and call it “cyan,” they see red and call it “cyan.” Accordingly, having the experience of red cannot be a matter of producing the right behavior. Our subjects suffering from an inverted spectrum *behave* around red just as we do, even calling it “red,” but actually experience a green-blue cyan color. According to behaviorism, the subjects are *experiencing* red; but this is false. Therefore, behaviorism is false.

There are other overwhelming arguments against behaviorism, but perhaps the biggest has been the realization from psychiatry, psychology, and physiology that events in the brain can explain behavior. If the relevant parts of the visual cortex are absent or damaged, for example, color vision can be affected, and our behavior will not be the same as the behavior of someone with a properly functioning visual cortex. During the two World Wars the evidence from neurology and from the hospitals mounted up. It began to look as though the state of the brain is what is making us behave in the way we do, or at the least allowing us to—though these are hardly the same thing. When in the 1950s the evidence for a causal explanation of behavior in the brain, or anyway a causal explanation of abnormal

behavior, became too telling, behaviorism started to lose almost all its popularity, and very quickly at that.

It was especially troubling that if, according to behaviorism, a mental state is a disposition to behave, then if what explains the behavior is the mental state, as we would ordinarily think, *we have to say that what explains the behavior is the disposition to behave in that way!* Thus behaviorism amounts to a tautology—a trivial truth—if there is such a thing as an explanation of the body's behavior by mental causes.

The Identity Theory

By the mid-1950s, when things began to change, they changed completely. Starting with a pioneering paper in 1956 by U. T. Place, more and more philosophers and scientists were persuaded that the explanation both of what people do and of what they experience lies in the brain. American and Australian philosophers in particular began to advance what became known as the “mind–brain identity theory,” or the “identity theory,” as it is called for short. This view, as its name suggests, is the claim that mind and brain, or anyway the relevant bits of the central nervous system, are identical, one and the same. Here too, the mind–body problem is solved at a stroke, by physicalism, by the denial that the mind is a nonphysical thing. Every

mental event is a physiological event within the nervous system. Accordingly, the theory that the mind is the brain has sometimes been known as “central-state” materialism, a materialism making the mind into the central nervous system, distinguishing it from the “peripheral-state” materialism of the behaviorists.

In its favor, the theory can be said to be commonsensical, given the facts of neurology such as the effects of brain damage, and it makes a great simplification in the philosophy of mind. But it is hardly an “astonishing hypothesis,” as Francis Crick claimed in a book of that title published in 1994. It is important and interesting, certainly, but not so astonishing. Like behaviorism, it solves the mind-body problem at a stroke, by denying that the mind is nonphysical. If this proposition about the mind is true, then the solution is, as before, impeccable. The mind is the brain and the brain is a physical thing, so the mind can interact with the rest of the body without difficulty. Yet we miss the essential thing needed for a solution: how has the physical, which has physical properties, turned into the mental, which has properties incompatible with being a part of the physical? What do neurons have when they fire that produces mind rather than electrical signals, or soap bubbles, for that matter?

Against the theory are also certain logical and philosophical difficulties. The central-state materialists do not claim and are bound not to claim that the word “mind”

means “brain,” which is fortunate for them, as “mind” as a matter of fact does not mean “brain.” If it did, the claim about the meanings of the words would make the main claim of central-state materialism (that the mind is the brain) into a necessary truth generated by the meanings of the two words. Its truth could have been discovered simply by looking in the dictionary. However, what the mind is was taken by the central-state materialists to be an empirical and factual question, not one of meaning. Central-state materialists, including Crick, took the question to be scientific, in just the same way as the question of what the gene or unity of heredity is was empirical and factual, to use the central-state materialists’ own favorite example. The gene turned out to be DNA, but this could not have been known from the meanings of words “gene” and “deoxyribonucleic acid.”

So far so good. But then there appeared an unpleasant proof from the world of logic. Identity, as it turns out, is always necessary. Suppose $a = b$. a has the following interesting property. It is necessarily identical with itself, a . Take this last statement, that a is necessarily identical with a . Substitute b for the second a ; we are entitled to do this, since we have supposed that $a = b$. But now it follows that a is necessarily identical with b . Accordingly, if central-state materialism is going to claim that the mind and the brain are not necessarily identical, it must itself be false. This proof was published by Saul Kripke in lectures given in

1970, and he developed extraordinarily interesting related arguments in the same work.³

Proofs of this sort, it should be noted, rely on the fact that the terms on either side of the identity sign, here “*a*” and “*b*,” are fixed names (“rigid designators,” as Kripke called them) and not descriptions that can be applied to different things. “The human mind” and “the human brain” are names, and so are “pain” and “events *a* in the thalamus, *b* the pre-frontal cortex, or *c* the primary and secondary somatosensory cortex (S1 and S2).”⁴ So the proof does not imply that it is somehow necessary that the Queen is Elizabeth II, which is true as I write. “Elizabeth II” is a name, but “the Queen,” even “the Queen of England” is really a compressed *description* that can apply to different persons, as it has done in the last hundred years. It is not a rigid designator because the place of the object of its description can be different objects.

Furthermore, the claim that the mind is the brain also turns out to be equivalent to the claim that the brain is the mind, since identity is what logicians and mathematicians call “commutative.” If $a = b$ then obviously $b = a$. But the claim that the brain is really at bottom the mind could hardly be expected to appeal to a hard-headed central-state *materialist*, since it makes a claim more suggestive of *idealism* (everything is mind) than of *materialism* (everything is matter).

What is a central-state materialist to do?

One answer was to take advantage of a distinction that had existed for some time in general philosophy, including metaphysics and the philosophy of art: the distinction between *types* and *tokens*. Take, for example, Edward Elgar's Cello Concerto in E minor. It has been played many times, including its disastrous premiere in 1919, Jacqueline du Pré's triumphant and elegiac performances in the 1960s, and hundreds of others. How many Elgar Cello Concertos are there? Could one say that there are hundreds? In that case, since Elgar wrote the work or works, he wrote hundreds of Cello Concertos. But he didn't. He was enormously hardworking, but not that hardworking. Or is there only one concerto? But then how could it appear in all sorts of different places and at all sorts of different times with so many different soloists? The answer developed by philosophers is that there is one concerto *type* and many concerto *tokens* or *instances*, in much the same way that there is one book called *Pride and Prejudice*, but many *copies* of the book. The copy both is and is not the work; it is a token of the work, but it is not the type. There is a difference between the Cello Concerto case and the case of the book, though, because there is nothing that could be regarded as the performance of *Pride and Prejudice*. But though what is played is "the music," as it is written, all the same it can be said that the glorious sound that is the Cello Concerto is not the sheet music, whereas the printed copies of the book are the novel.

The distinction has its difficulties, clearly, but it was used advantageously to distinguish two forms of the identity theory. There is the type *pain*, and there is the individual *pain* that is a token of the type. In the stronger and less plausible form of the theory it was the type or property *mental state* that was said to be identical with the type or property *brain state*. In the less sweeping and more convincing version, it was instead said to be just the one particular instance of a mental state that was identical with a particular instance of a brain state. It might be that two organisms both feel the same or a similar pain, but that they are not in the same brain state. They are in some brain state; and since it is implausible that everyone's physiological and psychological systems work in the same way, especially when we consider different organisms that have very different kinds of brains, it is much more plausible to identify *this* pain with *this* brain state, and accept the consequence that two individuals in the same psychological state may not be in the same physiological state. But they must be in *some* physiological state, with which the pain state is identical. So one is bound to wonder what makes all the tokens into tokens of the same type. Why are they all instances of pain?

In any case, it was suggested that the logical arguments against central-state materialism only worked against identities of types. That turned out not to be the case. The

arguments, as it was soon realized, worked equally well against identities of tokens.

Even before the logical proofs against central-state materialism were worked out and made public in the 1970s, however, it was already too late; central-state materialism was dead in the water. This came about not because of the intricate logical argumentation against it, but because a much more powerful view had arisen to take the place of central-state materialism, more in keeping with the science of the time.

Functionalism

The new view that took the place of central-state materialism was *functionalism*. It came upon the philosophical scene in 1967 with Hilary Putnam's "Psychological Predicates" and other subsequent papers.⁵ Putnam argues that pain is not a brain state, but another kind of state entirely. It is a state of a probabilistic automaton or a *Turing machine*. A Turing machine is in essence a computer, and it computes, having *computational* or *functional states* that are not its physical states. They are described completely differently, for one thing, and for another the computational states are not made of matter, but rather of a kind of functionality, if they can be said to be made of anything at all. One can also imagine that two Turing

machines could happen accidentally to be in just the same physical states, but in the process of performing different computations. So their computational states at that moment at which they are physically identical would not be the same states. So if mental states are computational states, as functionalism suggested, they are not the physical states of the organism.

The power of functionalism came from the interesting fact that it deployed to full effect the distinction between computer hardware and computer software. What is going on with functionalism is that the mind is compared to active software, not to rigid hardware. Even with ordinary computers, one can imagine that two laptops computing the same function, say, the multiplication 7×9 , might do it in very different physical ways. One might even consider an optical computer that does not work in the same way as an electronic computer, by electrons slowly pushing one another around through the different gates that make up the central processing unit. Clearly the two computers, optical and electronic, are not in the same physical state, since photons are not electrons. But the output (63) will always be the same given the same input (7×9). One can think of the function of the two machines as the same; for even their logical architecture might be quite different. Again, even two electronic computers might be running very different programs yet happen coincidentally at some instant to be in the same physical state.

Putnam had discovered the *multiple realizability thesis*, the proposition that one mental state can be realized in multiple and very different ways. Goats, birds, reptiles, and mollusks all feel pain, depending of course what your philosophy of animal minds is. But it is completely implausible to think that when they do, they are all in the precisely the same physiological brain state.

One might have thought, as Putnam pointed out, that the effect of the development of computers on the philosophy of mind was going to be materialistic, but in the event it was the reverse. The distinction between hardware and software allowed computing systems to be considered in abstraction from their physical states, and to highlight the difference between the computational or Turing-machine state, and the physical.

The time was right for functionalism, and it swept through the philosophy of mind in spite of some rear-guard action by central-state materialists. It rapidly became the preferred philosophy of mind of the artificial intelligentsia, those working in artificial intelligence, but also of many philosophers, especially philosophers of mind, and scientists in fields other than cognitive science.

How does functionalism solve the mind-body problem? The most obvious interpretation is that functionalism denies that the mind is a nonphysical thing, not because it takes the line that the mind is a physical thing, but because it takes the line that it is as wrong to think of

the mind just as a *thing*, as it would be to think of a computer program as just another physical thing. What is important about the program is not the physical states of the piece of tape or paper or electronic hardware as a thing, but its functionality.

Putnam has now recanted, arguing in 1991, against his own former functionalist self, that functionalism is false. One of his arguments is that *any* computational description of nonphysical properties can be applied to *any* physical thing, so that functionalism is completely trivial. This is interesting for us, however, whether it is right or not, because it suggests that, before he came to reject it, Putnam had taken functionalism to hold that physical systems, including mental ones, *do* have unique computational descriptions, and that this fact is what is behind the truth that the computational is not the physical. In that case, Putnam must have thought that functionalism solves the mind-body problem by denying the proposition that the mind is a nonphysical thing.

If this is right, it will come as no surprise that functionalism has a problem with qualia or phenomenal properties, just as behaviorism had had. One can easily conceive of two “systems” in just the same functional or computational or Turing-machine state, built into a robot with inputs having spectra inverted relative to one another, and computing on the basis of these inputs. Accordingly, the qualitative experience of the spectrum cannot be the same thing as

a functional or computational state. A Turing machine may be computing away without the slightest idea of what it is computing about, the color red, say. It computes, happily accepting inputs and giving outputs about colors, without having the slightest impression or idea what colors are.

An even more interesting idea is that functionalism is a form of property dualism, if it is taken to claim that it is false that the body is physical; for the body, including the brain, might be thought to have nonphysical or functional computational states.

Anomalous Monism

At about the same time as functionalism was changing the world of the philosophy of mind, the philosopher Donald Davidson was independently developing a deep and interesting view of the relationship between the mind and the physical world.

In a classic paper from 1970, "Mental Events," Davidson takes it as given that there are descriptions of events in the world, descriptions that are irreducibly mental, in the sense that they use mentalistic words that cannot be defined by physical terms, as well as physical events.⁶ So he subscribes to the essence of the propositions that the mind is nonphysical and that the body is physical, our (1) and

(2). He also takes the view that there are causal relations between at least some mental events and some physical events. But, he notes, such causal relations require a basis in a law that covers them. Where there is a causal relation, there is a law to cover it. There is no such thing as “singular” causation that works on one occasion but not on others. All this is easy enough to accept, until we reflect with Davidson that it is also the case that there are no strict laws covering the relations between the physical and mental events. There is no physical law that absolutely demands that when I decide to go to Italy to see my grandmother, I find my neurons firing in exactly this or that way.

Of course, one might doubt the general truth of this “anomalism of the mental,” because there are some pretty strict laws in psychophysics. An example is the so-called Weber–Fechner law for the perception of weight. The law states that in human perception there is a logarithmic relationship between the strength of the stimulus and the strength of response. A correspondingly greater increase in the stimulus is required to increase the same response.

In the first place, however, the law is actually not strict. It applies *moderately* well to human perception, but it only applies well over certain ranges of perception, such as the higher amplitudes in audition or hearing, and there are other limitations as well.

It might well be thought, of course, that though there are no strict psychophysical laws, this is hardly surprising,

because there are no strict physical laws either. Gravitational laws, for example, assume a hard vacuum, which never strictly exists. It should be conceded, though, that the laws of physics and chemistry are very much stricter than the laws of psychophysics.

We should also note that Davidson's real interest was in paradigmatically *psychological* laws as they apply to human behavior, or the more rational and conscious parts of human behavior, such as my decision to go to visit my grandmother in Italy, not in perception. There really is no law about such an event or about an event "so described," as Davidson puts it. The intention concerns the rational end of human behavior, and rationality could hardly be codified in such a way as to connect up with the world of scientific law. But our concepts of the mental are tied up with rationality, for example in such ideas as "reasonable," "intent," "intention," "thoughtfulness," and so on.

For Davidson there are causal relations between the mental and the physical, and causal relations demand strict laws, but there are no strict laws between the physical and the mental. And here we have a huge and fascinating problem. We have an inconsistent *triad*, one indeed that has a definite relationship to our original inconsistent tetrad. If the mental and physical interact, and causal relations demand strict laws, then there certainly ought to be strict laws governing physical and mental events. But

there aren't, according to Davidson, and we have a contradiction on our hands.

What is actually happening is that Davidson is initially affirming that physical and nonphysical things cannot interact, because that would require strict causal laws between the mental and the physical. He then notes that mind and body *do* interact, but, as he finally puts it, they can only do so under a nonmental *vocabulary*, one that is not physical in nature. What is left is that mental descriptions are anomalous, in that they do not connect systematically with scientific explanations. It is perhaps worth noting that Davidson began his academic career with a PhD in classical Greek philosophy, and that he has always been alive to the richness and variety of language about the mind.

There are things other than mental events that have anomalous descriptions. One might take an interest in things that are *cheap*, for example, without thinking that cheap things have anything in common that could relate them to the physical world by means of strict laws. There are no strict laws of cheapness, if you like. "Cheap" is a vague, idiosyncratic, and interest-relative predicate. It reflects our everyday behavior and practices in such a way that it could never become a word used in a strict science, even economics. And so it is with mental words. They reflect our rational interests, for example in explaining

everyday actions, but not within the framework of physical law.

By the way, Davidson also has an argument here for the conclusion that mental events must be token-identical with certain physical events. Since mental events, so described, do not fall under strict laws, and since they do interact with physical events, they must fall under physical laws. Hence they must be physical events. But they must also not be physical events *described* as such, and so they are not type-identical with physical events. So they are token-identical with physical events. This is Davidson's argument that every mental event is actually some physical event. It is certainly a brilliant line of thought.

In his overall argument concerning the mind-body relation, Davidson can be taken to be arguing that:

- (1) The mind is a nonphysical thing

(in the sense that *descriptions* of the mind are couched in nonphysical terms, but not in the sense that it is not a physical object).

- (2) The body is physical.
- (3) The mind and the body interact.
- (4) Physical and nonphysical things cannot interact

(since (4) would require strict causal laws between them, and there are none).

The distinctively Davidsonian catch that we see in the ambiguity of (1) is that under another and peculiarly human descriptive vocabulary, mental descriptions *are* non-physical, not in the sense that these descriptions are not, say, written in physical ink or spoken in physical words, but in the sense that they do not use any of the words or symbols of physics, and do use other “mental” or psychological words. Mental events can in one sense truly be said to be physical, but they can also be described in a nonphysical vocabulary, just as objects can in one sense truly be said to be physical, but they can also be described in the nonphysical vocabulary of home economics. “That’s a cheap bag of tomatoes—let’s buy it,” we might say. That doesn’t mean that the tomatoes are not physical things.

For all its undoubted charm, we should not allow Davidson’s view to cause us to forget the logical difficulties with central-state materialism, of the type or token variety, the powerful insights of functionalism, nor the difficulty that any form of materialism has dealing with qualia. When we allude to qualia, to colors for example, we are not just adopting a funny new vocabulary, in addition to talk about electromagnetic radiation, that happens to suit us in our dealings with the world. The new vocabulary is not just a different way of talking about electromagnetic radiation.

It is a way of talking about something completely different: colors. Colors have properties that are not strictly physical—for example, brightness. Brightness is not physical. It is related to *luminance*, the narrowly physical and physically defined concept, which is about how much radiation is transmitted, emitted, or reflected by a particular unit area. When we say that yellow is a bright color, this has an entirely nonphysical meaning, one that can in principle be determined and can only be determined by direct observation, without the measurement of luminance of the yellow colored area. *Brightness* is not a concept to be found in physics but, on the standard view, a concept to be found in psychology.

Eliminativism

With anomalous monism, one has the feeling that the mental has been spirited away, as some, not including Davidson, might think it deserves to be. Perhaps it would be better for Davidson to allow that that there is no such *thing* as the mental, though there are mental *vocabularies*, descriptions, explanations and ways of talking, or mental concepts, but then one starts to worry that the mental is being swept under a convenient linguistic rug.

With the philosophy of mind known as “eliminativism” or “eliminative materialism,” we have the straight recognition that talk about the mental will not fit into the scheme of things given to us by the study of the physical world as it applies to human beings, or to any other part of science. Eliminativists admit that mental concepts and terms cannot be reduced to scientific physiological ones. They draw the conclusion that in a completed neuroscience there is no need and no room for mental terms and concepts, and that statements about things mental are just false. These statements are relics of an outmoded psychology and psychophysics, just as statements about witches are relics of outmoded an outmoded view of human nature. There are no witches, and witches are a product of superstition. Nor can we relate witches to concurrent physical events, such as pot-stirring and the training of cats by night. Similarly, says the eliminativist, there are no hopes, fears, beliefs, and desires; they are a product of an inherited form of language that has no basis in science, explains nothing, and has no use beyond the parochial view that belongs in the gossip-filled village shop, and certainly has no use in a scientific laboratory.

Unlike Davidson, eliminative materialists, of whom the most distinguished are Paul and Patricia Churchland and Stephen Stich, take the view that the sentences of the psychology of everyday life that refer to hopes, fears, beliefs, and desires are a sort of a theory, but a completely

Eliminativists admit
that mental concepts
and terms cannot be
reduced to scientific
physiological ones.

false one. What is called “folk psychology” by its detractors, on the analogy with “folk remedies,” “folklore,” and so on, is false:

The common-sense conception of psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by a completed neuroscience.⁷

For Davidson, on the other hand, folk psychology is not explanatory, and it is not a theory at all. That role is reserved for physics. But it is descriptive.

Folk psychology, writes Paul Churchland, “suffers explanatory failures on an epic scale, ... has been stagnant for at least twenty-five centuries, and ... 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.”⁸ Most philosophers disagree with Churchland that there is something called folk psychology which is a *theory* that makes predictions, the so-called *theory theory*. It is rather a loose set of *concepts* that we employ in ordinary life, concepts like *family*, in the social world, or the *state*, in the political world, or *work of art* in the world of art. Propositions using such concepts are not radically false. And nor should

such concepts be swept away in return for some supposedly more profound and accurate scientific concepts, any more than those of cooking or of personal relationships.

There is another problem with Churchland's claim. Arithmetic, for example, has stagnated for far longer than folk psychology, if that means merely that it has not changed. There have been no changes in elementary arithmetic since it was discovered. Multiplication, division, addition, and subtraction—all have “stagnated.” Projective geometry, to take another more high-powered example, is in essence complete, and has been since the late nineteenth century. Why then should the concepts of the psychology of ordinary life also not remain undisturbed?

The answer, writes Churchland, is that folk psychology should be displaced because it has not explained mental illness (all of it?), creative imagination, individual differences in intelligence, sleep, the ability to hit targets with projectiles such as baseballs, 3D perception, all the visual illusions, memory, the speed of memory, learning (including learning in prelinguistic infants), and so on and so forth. On all these, folk psychology sheds “negligent light.”

The list sets a high bar indeed—too high. For *science* has not explained what sleep is, nor what mental illness is (which incidentally it could not do on the eliminativist view, since “mental illness” is a folk psychological concept, and so it must be “radically false” that people have mental illnesses), nor what creative imagination is (*imagination* is

another folk psychological concept, however), and so on. Perhaps there is some conviction that *only* science could explain all these mysterious things, but that is the *conclusion* for which Churchland is arguing, not a *premise* from which he is entitled to argue.

And consider. Elementary arithmetic has failed dismally in the last two thousand five hundred years, and indeed the whole of mathematics has failed, to determine whether there is an odd perfect number, the truth of the Goldbach conjecture, the solution to the Collatz problem, the twin prime conjecture, and so on. Arithmetic has really stagnated, no? And it should pull up its socks. Perhaps we should replace it with neuroscience, which has solved all sorts of important problems.

It is again very obvious how eliminative materialism solves the mind-body problem. Proposition (1), that the mind is a nonphysical thing, is false, not because the mind *is* physical, but because there is no mind. Nothing, including the mind, is nonphysical. The existence of something called “the mind,” and all its works, is part of a “radically false” folk mythology.

Again, we have a completely successful solution to the mind-body problem, and again we have a view that is itself every bit as hard to believe as the mind-body problem is said to be intractable. The clear success of a solution seems to stand in inverse relationship to its believability.