

Matter and Consciousness

Paul Churchland, 1984

Chapter 2: The Ontological Problem (the Mind-Body Problem)

5. *Eliminative Materialism*

The identity theory was called into doubt not because the prospects for a materialist account of our mental capacities were thought to be poor, but because it seemed unlikely that the arrival of an adequate materialist theory would bring with it the nice one-to-one match-ups, between the concepts of folk psychology and the concepts of theoretical neuroscience, that intertheoretic reduction requires. The reason for that doubt was the great variety of quite different physical systems that could instantiate the required functional organization. *Eliminative materialism* also doubts that the correct neuroscientific account of human capacities will produce a neat reduction of our common-sense framework, but here the doubts arise from a quite different source.

As the eliminative materialists see it, the one-to-one match-ups will not be found, and our common-sense psychological framework will not enjoy an intertheoretic reduction, *because our common-sense psychological framework is a false and radically misleading conception of the causes of human behavior and the nature of cognitive activity*. On this view, folk psychology is not just an incomplete representation of our inner natures; it is an outright *misrepresentation* of our internal states and activities. Consequently, we cannot expect a truly adequate neuroscientific account of our inner lives to provide theoretical categories that match up nicely with the categories of our commonsense framework. Accordingly, we must expect that the older framework will simply be eliminated, rather than be reduced, by a matured neuroscience.

Historical Parallels

As the identity theorist can point to historical cases of successful intertheoretic reduction, so the eliminative materialist can point to historical cases of the outright elimination of the ontology of an older theory in favor of the ontology of a new and superior theory. For most of the eighteenth and nineteenth centuries, learned people believed that heat was a subtle *fluid* held in bodies, much in the way water is held in a sponge. A fair body of moderately successful theory described the way this fluid substance—called “caloric”—flowed within a body, or from one body to another, and how it produced thermal expansion, melting, boiling, and so forth. But by the end of the last century it had become abundantly clear that heat was not a substance at all, but just the energy of motion of the trillions of jostling molecules that makeup the heated body itself. The new theory—the “corpuscular/kinetic theory of matter and heat”—was much more successful than the old in explaining and predicting the thermal behavior of bodies. And since we were unable to *identify* caloric fluid with kinetic energy (according to the old theory, caloric is a material *substance*; according to the new theory, kinetic energy is a form of *motion*), it was finally agreed that there is *no such thing* as caloric. Caloric was simply eliminated from our accepted ontology.

A second example. It used to be thought that when a piece of wood burns, or a piece of metal rusts, a spiritlike substance called “phlogiston” was being released: briskly, in the former case, slowly in the latter. Once gone, that “noble” substance left only a base pile of ash or rust. It later came to be appreciated that both processes involve, not the loss of something, but the *gaining* of a substance taken from the atmosphere: oxygen. Phlogiston emerged, not as an incomplete description of what was going on, but as a radical misdescription. Phlogiston was therefore not suitable for reduction to or identification with some notion from within the new oxygen chemistry, and it was simply eliminated from science.

Admittedly, both of these examples concern the elimination of something nonobservable, but our history also includes the elimination of certain widely accepted “observables.” Before Copernicus’ views became

available, almost any human who ventured out at night could look up at *the starry sphere of the heavens*, and if he stayed for more than a few minutes he could also see that it *turned*, around an axis through Polaris. What the sphere was made of (crystal?) and what made it turn (the gods?) were theoretical questions that exercised us for over two millennia. But hardly anyone doubted the existence of what everyone could observe with their own eyes. In the end, however, we learned to reinterpret our visual experience of the night sky within a very different conceptual framework, and the turning sphere evaporated.

Witches provide another example. Psychosis is a fairly common affliction among humans, and in earlier centuries its victims were standardly seen as cases of demonic possession, as instances of Satan's spirit itself, glaring malevolently out at us from behind the victims' eyes. That witches exist was not a matter of any controversy. One would occasionally see them, in any city or hamlet, engaged in incoherent, paranoid, or even murderous behavior. But observable or not, we eventually decided that witches simply do not exist. We concluded that the concept of a witch is an element in a conceptual framework that misrepresents so badly the phenomena to which it was standardly applied that literal application of the notion should be permanently withdrawn. Modern theories of mental dysfunction led to the elimination of witches from our serious ontology.

The concepts of folk psychology—belief, desire, fear, sensation, pain, joy, and so on—await a similar fate, according to the view at issue. And when neuroscience has matured to the point where the poverty of our current conceptions is apparent to everyone, and the superiority of the new framework is established, we shall then be able to set about reconceiving our internal states and activities, within a truly adequate conceptual framework at last. Our explanations of one another's behavior will appeal to such things as our neuropharmacological states, the neural activity in specialized anatomical areas, and whatever other states are deemed relevant by the new theory. Our private introspection will also be transformed, and may be profoundly enhanced by reason of the more accurate and penetrating framework it will have to work with—just as the astronomer's perception of the night sky is much enhanced by the

detailed knowledge of modern astronomical theory that he or she possesses.

The magnitude of the conceptual revolution here suggested should not be minimized: it would be enormous. And the benefits to humanity might be equally great. If each of us possessed an accurate neuroscientific understanding of (what we now conceive dimly as) the varieties and causes of mental illness, the factors involved in learning, the neural basis of emotions, intelligence, and socialization, then the sum total of human misery might be much reduced. The simple increase in mutual understanding that the new framework made possible could contribute substantially toward a more peaceful and humane society. Of course, there would be dangers as well: increased knowledge means increased power, and power can always be misused.

Arguments for Eliminative Materialism

The arguments for eliminative materialism are diffuse and less than decisive, but they are stronger than is widely supposed. The distinguishing feature of this position is its denial that a smooth intertheoretic reduction is to be expected—even a species-specific reduction—of the framework of folk psychology to the framework of a matured neuroscience. The reason for this denial is the eliminative materialist's conviction that folk psychology is a hopelessly primitive and deeply confused conception of our internal activities. But why this low opinion of our common-sense conceptions?

There are at least three reasons. **First**, the eliminative materialist will point to the widespread explanatory, predictive, and manipulative failures of folk psychology. So much of what is central and familiar to us remains a complete mystery from within folk psychology. We do not know what *sleep* is, or why we have to have it, despite spending a full third of our lives in that condition. (The answer, "For rest," is mistaken. Even if people are allowed to rest continuously, their need for sleep is undiminished. Apparently, sleep serves some deeper functions, but we do not yet know what they are.) We do not understand how *learning* transforms each of us from a gaping infant to a cunning adult, or how

differences in *intelligence* are grounded. We have not the slightest idea how *memory* works, or how we manage to retrieve relevant bits of information instantly from the awesome mass we have stored. We do not know what *mental illness* is, nor how to cure it.

In sum, the most central things about us remain almost entirely mysterious from within folk psychology. And the defects noted cannot be blamed on inadequate time allowed for their correction, for folk psychology has enjoyed no significant changes or advances in well over 2,000 years, despite its manifest failures. Truly successful theories may be expected to reduce, but significantly unsuccessful theories merit no such expectation.

This **argument from explanatory poverty** has a further aspect. So long as one sticks to normal brains, the poverty of folk psychology is perhaps not strikingly evident. But as soon as one examines the many perplexing behavioral and cognitive deficits suffered by people with *damaged* brains, one's descriptive and explanatory resources start to claw the air . . . As with other humble theories asked to operate successfully in unexplored extensions of their old domain (for example, Newtonian mechanics in the domain of velocities close to the velocity of light, and the classical gas law in the domain of high pressures or temperatures), the descriptive and explanatory inadequacies of folk psychology become starkly evident.

The **second** argument tries to draw an inductive lesson from our conceptual history. Our early folk theories of motion were profoundly confused, and were eventually displaced entirely by more sophisticated theories. Our early folk theories of the structure and activity of the heavens were wildly off the mark, and survive only as historical lessons in how wrong we can be. Our folk theories of the nature of fire, and the nature of life, were similarly cockeyed. And one could go on, since the vast majority of our past folk conceptions have been similarly exploded. All except folk psychology, which survives to this day and has only recently begun to feel pressure. But the phenomenon of conscious intelligence is surely a more complex and difficult phenomenon than any of those just listed. So far as accurate understanding is concerned, it would be a *miracle* if we had got *that* one right the very first time, when

we fell down so badly on all the others. Folk psychology has survived for so very long, presumably, not because it is basically correct in its representations, but because the phenomena addressed are so surpassingly difficult that any useful handle on them, no matter how feeble, is unlikely to be displaced in a hurry.

A **third** argument attempts to find an a priori advantage for eliminative materialism over the identity theory and functionalism. It attempts to counter the common intuition that eliminative materialism is distantly possible, perhaps, but is much less probable than either the identity theory or functionalism. The focus again is on whether the concepts of folk psychology will find vindicating match-ups in a matured neuroscience. The eliminativist bets no; the other two bet yes. (Even the functionalist bets yes, but expects the match-ups to be only species-specific, or only person-specific. Functionalism, recall, denies the existence only of *universal* type/type identities.)

The eliminativist will point out that the requirements on a reduction are rather demanding. The new theory must entail a set of principles and embedded concepts that mirrors very closely the specific conceptual structure to be reduced. And the fact is, there are vastly many more ways of being an explanatorily successful neuroscience while *not* mirroring the structure of folk psychology, than there are ways of being an explanatorily successful neuroscience while also *mirroring* the very specific structure of folk psychology. Accordingly, the a priori probability of eliminative materialism is not lower, but substantially *higher* than that of either of its competitors. One's initial intuitions here are simply mistaken.

Granted, this initial a priori advantage could be reduced if there were a very strong presumption in favor of the truth of folk psychology—true theories are better bets to win reduction. But according to the first two arguments, the presumptions on this point should run in precisely the opposite direction.

Arguments Against Eliminative Materialism

The initial plausibility of this rather radical view is low for almost everyone, since it denies deeply entrenched assumptions. That is at best a question begging complaint, of course, since those assumptions are precisely what is at issue. But the following line of thought does attempt to mount a real argument.

Eliminative materialism is false, runs the argument, because one's introspection reveals directly the existence of pains, beliefs, desires, fears, and so forth. Their existence is as obvious as anything could be.

The eliminative materialist will reply that this argument makes the same mistake that an ancient or medieval person would be making if he insisted that he could just see with his own eyes that the heavens form a turning sphere, or that witches exist. The fact is, all observation occurs within some system of concepts, and our observation judgments are only as good as the conceptual framework in which they are expressed. In all three cases—the starry sphere, witches, and the familiar mental states—precisely what is challenged is the integrity of the background conceptual frameworks in which the observation judgments are expressed. To insist on the validity of one's experiences, *traditionally interpreted*, is therefore to beg the very question at issue. For in all three cases, the question is whether we should reconceive the nature of some familiar observational domain.

A second criticism attempts to find an incoherence in the eliminative materialist's position. The bald statement of eliminative materialism is that the familiar mental states do not really exist. But that statement is meaningful, runs the argument, only if it is the expression of a certain *belief*, and an *intention* to communicate, and a *knowledge* of the language, and so forth. But if the statement is true, then no such mental states exist, and the statement is therefore a meaningless string of marks or noises, and cannot be true. Evidently, the assumption that eliminative materialism is true entails that it cannot be true.

The hole in this argument is the premise concerning the conditions necessary for a statement to be meaningful. It begs the question. If eliminative materialism is true, then meaningfulness must have some

different source. To insist on the “old” source is to insist on the validity of the very framework at issue. Again, an historical parallel may be helpful here. Consider the medieval theory that being biologically *alive* is a matter of being ensouled by an immaterial *vital spirit*. And consider the following response to someone who has expressed disbelief in that theory.

My learned friend has stated that there is no such thing as vital spirit. But this statement is incoherent. For if it is true, then my friend does not have vital spirit, and must therefore be *dead*. But if he is dead, then his statement is just a string of noises, devoid of meaning or truth. Evidently, the assumption that antivitalism is true entails that it cannot be true! Q.E.D.

This second argument is now a joke, but the first argument begs the question in exactly the same way.

A final criticism draws a much weaker conclusion, but makes a rather stronger case. Eliminative materialism, it has been said, is making mountains out of molehills. It exaggerates the defects in folk psychology, and underplays its real successes. Perhaps the arrival of a matured neuroscience will require the elimination of the occasional folk-psychological concept, continues the criticism, and a minor adjustment in certain folk-psychological principles may have to be endured. But the large-scale elimination forecast by the eliminative materialist is just an alarmist worry or a romantic enthusiasm.

Perhaps this complaint is correct. And perhaps it is merely complacent. Whichever, it does bring out the important point that we do not confront two simple and mutually exclusive possibilities here: pure reduction versus pure elimination. Rather, these are the endpoints of a smooth spectrum of possible outcomes, between which there are mixed cases of partial elimination and partial reduction. Only empirical research... can tell us where on that spectrum our own case will fall. Perhaps we should speak here, more liberally, of “revisionary materialism,” instead of concentrating on the more radical possibility of an across-the-board elimination. Perhaps we should. But it has been my aim in this section to make it at least intelligible to you that our collective conceptual destiny lies substantially toward the revolutionary end of the spectrum.