I

identity theories What is the relation between mind and physical reality? Well-established schools of thought give starkly opposing answers to this question. Descartes insisted that mental phenomena are non-physical in nature. This view seems inviting because mental phenomena are indisputably different from everything else. Moreover, it's safe to assume that all phenomena that aren't mental have some physical nature. So it may seem that the best way to explain how the mental differs from everything else is to hypothesize that mind is non-physical in nature.

But that hypothesis is not the only way to explain how mind differs from everything else. It's also possible that mental phenomena are instead just a special case of physical phenomena; they would then have properties that no other physical phenomena have, but would still themselves be physical. This explanation requires that we specify what is special about mental phenomena which makes them different from everything else. But we must specify that in any case, just in order to understand the nature of the mental. Characterizing mental phenomena negatively, simply as not being physical, does little to help us understand what it is for something to be mental.

The claim that mental phenomena are a special kind of physical phenomenon is the root idea of mind–body materialism (also called PHYSICALISM). And the version of that thesis that's been most widely defended over the last three decades is the identity theory of mind.

In Descartes' time the issue between materialists and their opponents was framed in terms of substances. Materialists such as Thomas Hobbes and Pierre Gassendi maintained that people are physical systems with abilities that no other physical systems have; people, therefore, are special kinds of physical substance. Descartes' DUALISM, by contrast, claimed that people consist of two distinct substances that interact causally: a physical body and a non-physical unextended substance. The traditional conception of substance, however, introduces extraneous issues, which have no bearing on whether mental phenomena are physical or non-physical. And in any case, even those who agree with Descartes that the mental is non-physical have today given up the idea that there are non-physical substances. It's now widely accepted on all sides that people are physical organisms with two distinctive kinds of states: physical states such as standing and walking, and mental states such as thinking and feeling.

Accordingly, the issue of whether the mental is physical or non-physical is no longer cast in terms of whether people, and other creatures that have the ability to think and sense, are physical or non-physical substances. Rather, that question is put in terms of whether the distinctively mental states of thinking, sensing, and feeling are physical states or non-physical states. The identity theory is the materialist thesis that every mental state is physical, that is, that every mental state is identical with some physical state.

PROPERTIES OF MENTAL STATES

If mental states are identical with physical states, presumably the relevant physical states are various sorts of neural states. Our concepts of mental states such as thinking, sensing, and feeling are of course different from our concepts of neural states, of whatever sort. But that's no problem for the

identity theory. As J. J. C. Smart (1962), who first argued for the identity theory, emphasized, the requisite identities don't depend on our concepts of mental states or the meanings of mental terms. For *a* to be identical with *b*, *a* and *b* must have exactly the same properties, but the terms '*a*' and '*b*' need not mean the same (*see* LEIBNIZ'S LAW).

But a problem does seem to arise about the properties of mental states. Suppose pain is identical with a certain firing of c-fibres. Although a particular pain is the very same state as a neural firing, we identify that state in two different ways: as a pain and as a neural firing. The state will therefore have certain properties in virtue of which we identify it as a pain and others in virtue of which we identify it as a neural firing. The properties in virtue of which we identify it as a pain will be mental properties, whereas those in virtue of which we identify it as a neural firing will be physical properties. This has seemed to many to lead to a kind of dualism at the level of the properties of mental states. Even if we reject a dualism of substances and take people simply to be physical organisms, those organisms still have both mental and physical states. Similarly, even if we identify those mental states with certain physical states, those states will nonetheless have both mental and physical properties. So disallowing dualism with respect to substances and their states simply leads to its reappearance at the level of the properties of those states.

There are two broad categories of mental property. Mental states such as thoughts and desires, often called PROPOSITIONAL ATTITUDES, have CONTENT that can be described by 'that' clauses. For example, one can have a thought, or desire, that it will rain. These states are said to have intentional properties, or INTENTIONALITY. SENSATIONS, such as pains and sense impressions, lack intentional content, and have instead qualitative properties of various sorts.

The problem just sketched about mental properties is widely thought to be most pressing for sensations, since the painful quality of pains and the red quality of visual sensations seem to be irretrievably non-physical. So even if mental states are all identical with physical states, these states appear to have properties that aren't physical. And if mental states do actually have non-physical properties, the identity of mental with physical states won't sustain a thoroughgoing mind–body materialism.

Smart's reply to this challenge is that, despite initial appearances, the distinctive properties of sensations are neutral as between being mental or physical; in the term Smart borrowed from Gilbert RYLE, they are topic neutral. My having a sensation of red consists in my being in a state that's similar, in respects that we need not specify, to something that occurs in me when I'm in the presence of certain stimuli. Because the respect of similarity isn't specified, the property is neither distinctively mental nor distinctively physical. But everything is similar to everything else in some respect or other. So leaving the respect of similarity unspecified makes this account too weak to capture the distinguishing properties of sensations.

A more sophisticated reply to the difficulty about mental properties is due independently to D. M. Armstrong (1968) and David LEWIS (1972), who argue that for a state to be a particular sort of intentional state or sensation is for that state to bear characteristic causal relations to other particular occurrences. The properties in virtue of which we identify states as thoughts or sensations will still be neutral as between being mental or physical, since anything can bear a causal relation to anything else. But causal connections have a better chance than similarity in some unspecified respect of capturing the distinguishing properties of sensations and thoughts.

This causal theory is appealing (see the following section). Still, it's misguided to try to construe the distinctive properties of mental states as being neutral as between being mental or physical. To be neutral as regards being mental or physical is to be neither distinctively mental nor distinctively physical. But since thoughts and sensations

are distinctively mental states, for a state to be a thought or a sensation is perforce for it to have some characteristically mental property. We inevitably lose the distinctively mental if we construe these properties as being neither mental nor physical.

Not only is the topic-neutral construal misguided; the problem it was designed to solve is equally so. That problem stemmed from the idea that the mental must have some non-physical aspect, if not at the level of people or their mental states, then at the level of the distinctively mental properties of those states (*see* PROPERTY).

But the idea that the mental is in some respect non-physical cannot be assumed without argument. Plainly, the distinctively mental properties of mental states are unlike any other properties we know about. Only mental states have properties that are at all like the qualitative properties of sensations. And arguably nothing but mental states have properties that are anything like the intentional properties of THOUGHTS and DESIRES. But this doesn't show that these mental properties are not physical properties. Not all physical properties are like the standard cases; so mental properties might still be special kinds of physical properties. Indeed, it's question begging to assume otherwise. The doctrine that the mental properties of mental states are non-physical properties is simply an expression of the CARTESIAN doctrine that the mental is automatically non-physical.

To settle whether or not those mental properties are non-physical, we would need a positive account of what those properties are. Proposals are available that would account for intentional properties wholly in physical terms (*see* DENNETT, DRETSKE, and FODOR), and perhaps one of these will prove correct. It's been more difficult to give a positive account of the qualitative properties of sensations, and that's led some to conclude that such properties will inevitably turn out to be non-physical. But it's plainly unfounded to infer from the difficulty in explaining something to its being non-physical.

It's sometimes held that properties should count as physical properties only if they can be defined using the terms of physics. This is far too restrictive. Nobody would hold that to reduce biology to physics, for example, we must define all biological properties using only terms that occur in physics. And even putting REDUCTION aside, if certain biological properties couldn't be so defined, that wouldn't mean that those properties were in any way non-physical. The sense of 'physical' that's relevant here must be broad enough to include not only biological properties, but also most commonsense, macroscopic properties. Bodily states are uncontroversially physical in the relevant way. So we can recast the identity theory as asserting that mental states are identical with bodily states.

TYPES AND TOKENS

There are two ways to take the claim that every mental state is identical with some bodily state. It might mean identity at the level of types, that is, that every mental-state type is identical with some physical-state type. Such type identity would hold if all the instances of a particular type of mental state are also instances of a particular type of bodily state. This is called the type identity theory.

But the identity claim might instead mean only that every instance of a mental state is identical with an instance of a bodily state, of some type or other. On this construal, the various types of mental state wouldn't have to correspond to types of bodily state; instances of a single mental type might be identical with tokens of distinct bodily types. This weaker claim is known as the token identity theory.

There's reason to doubt that the type identity theory is true. It's plausible that organisms of different species may share at least some types of mental state – say, pain – even if their anatomical and physiological differences are so great that they can't share the relevant types of bodily state. No single bodily-state type would then correspond to these mental-state types. This possibility is called the multiple realizability of mental

states. It's conceivable, of course, that biology will someday type physiological states in a way that corresponds tolerably well with types of mental state, but we can have no guarantee that this will happen.

Even if no physiological types correspond to types of mental state, the causal theory of Lewis and Armstrong would allow us to identify types of mental state with types described in other terms. On Lewis's version of the theory, mental states are whatever states occupy the causal roles specified by all our commonsense psychological platitudes, taken together. The various types of mental state correspond to the various causal roles thus specified; mental-state tokens are of a particular mental type if they occupy the causal role that defines that type. These causal roles involve causal ties to behaviour and stimuli and to other states that occupy these causal roles. Such a theory, which defines mental-state types in terms of causal roles, is often called FUNCTIONALISM.

One could imagine that the individual states that occupy the relevant causal roles turn out not to be bodily states; for example they might instead be states of an Cartesian unextended substance. But it's overwhelmingly likely that the states that do occupy those causal roles are all tokens of bodily-state types. So the causal theory, together with this empirical likelihood, sustains at least the token identity theory. Moreover, this version of the causal theory bypasses the problematic idea that the mental properties of those states are neutral as between being mental or physical, since mental-state types are determined by our psychological platitudes.

To defend the type identity theory as well, however, would require showing that all mental-state tokens that occupy a particular causal role also fall under a single physiological type. Lewis (1980) expects substantial uniformity of physiological type across the tokens of each mental-state type, at least within particular populations of creatures. But if tokens of different physiological types do occupy the same causal role, that would undermine the type identity theory, or at least make it relative to certain populations.

Multiple realizability is the possibility that mental-state types are instantiated by states of distinct physiological types. It's an empirical matter whether that's actually the case. If it is, physical-state types don't correspond to mental-state types, and the type identity theory is false.

But one might, with Hilary PUTNAM (1975), construe the type identity theory more strongly, as claiming that the mental properties that define the various types of mental state are identical with physical properties. And that's false even if the tokens of each mental-state type fall under a single physiological type; the property of occupying a particular causal role is plainly not identical with the property of belonging to a particular physiological type. On this construal, no empirical findings are needed to refute the type identity theory. But it's more reasonable to construe the type identity theory less strongly, as requiring the claim only that all tokens of a particular mental-state type fall under a single physiological type.

Donald DAVIDSON (1970) has used different considerations to argue that mental-state types correspond to no physiological types, but that the token identity theory is nonetheless correct. Plainly, mental and bodily events cause each other. Moreover, as Davidson reasonably holds, one event token can cause another only if that causal connection instantiates some explanatory law. But Davidson also insists that an event token belongs to a particular mental type only relative to certain background assumptions about meaning and RATIONALITY. Tokens of physical events, by contrast, belong to whatever physical type they do independently of any such background assumptions. Davidson infers that there can be no strict laws connecting physical and mental events. But if so, how can mental and bodily events cause each other? (*See* REASONS AND CAUSES.)

Davidson's solution relies on the fact that explanatory laws describe events in particular ways and a different description of the

same events might not sustain the explanatory connection. So the impossibility of laws connecting mental and physical events means only that no laws can connect physical events, described as such, with mental events, described as such. To interact causally, events must figure in explanatory laws. So each mental-event token that interacts causally with a bodily event can figure in a law only if that mental-event token can also be described in purely physical terms. The considerations that preclude laws connecting mental with physical events presumably show also that no physical types correspond to any mental-state types. But since we can describe every mental-event token in physical terms, that token will be identical with some physical-event token. This intriguing argument is difficult to evaluate, mainly because it's unclear exactly why background assumptions about meaning and rationality should preclude laws connecting events described in mental terms with those described physically.

In order for causal interactions between mental and bodily events to fall under laws that describe events solely in physical terms, physically indistinguishable events must be mentally indistinguishable, though not necessarily the other way around. That relationship is known as SUPERVENIENCE; in this case, mental properties would be said to *supervene* on physical properties. Jaegwon Kim (1984) has usefully explored such supervenience as a way to capture the relation between mental and physical.

ELIMINATIVE MATERIALISM

The Cartesian doctrine that the mental is in some way non-physical is so pervasive that even advocates of the identity theory have sometimes accepted it, at least tacitly. The idea that the mental is non-physical underlies, for example, the insistence by some identity theorists that mental properties are really neutral as between being mental or physical. To be neutral in this way, a property would have to be neutral as to whether it's mental at all. Only if one thought that being mental meant being non-physical would one hold that defending materialism required showing that ostensibly mental properties are neutral as regards whether or not they're mental.

But holding that mental properties are non-physical has a cost that is usually not noticed. A phenomenon is mental only if it has some distinctively mental property. So, strictly speaking, a materialist who claims that mental properties are non-physical would have to conclude that no mental phenomena exist. This is the ELIMINATIVE-MATERIALIST position advanced by Richard Rorty (1979). (*See An Essay on Mind* section 3.7; ELIMINATIVISM.)

According to Rorty, 'mental' and 'physical' are incompatible terms. Nothing can be both mental and physical; so mental states cannot be identical with bodily states. Rorty traces this incompatibly to our views about incorrigibility; 'mental' and 'physical' are incompatible terms because we regard as incorrigible reports of one's own mental states, but not reports of physical occurrences. But he also argues that we can imagine a people who describe themselves and each other using terms just like our mental vocabulary, except that those people don't take the reports made with that vocabulary to be incorrigible. Since Rorty takes a state to be a mental state only if one's reports about it are taken to be incorrigible, his imaginary people don't ascribe mental states to themselves or each other. But the only difference between their language and ours is that we take as incorrigible certain reports which they don't. So their language has no less descriptive or explanatory power than ours. Rorty concludes that our mental vocabulary is idle, and that there are no distinctively mental phenomena.

This argument hinges on building incorrigibility into the meaning of the term 'mental'. If we don't, the way is open to interpret Rorty's imaginary people as simply having a different theory of mind from ours, on which reports of one's own mental states aren't incorrigible. Their reports would thus be about mental states, as construed by their theory. Rorty's thought experiment would then provide reason to conclude not

that our mental terminology is idle, but only that this alternative theory of mental phenomena is correct. His thought experiment would thus sustain the non-eliminativist view that mental states are bodily states. Whether Rorty's argument supports his eliminativist conclusion or the standard identity theory, therefore, depends solely on whether or not one holds that the mental is in some way non-physical.

Paul M. Churchland (1981) advances a different argument for eliminative materialism. According to Churchland, the commonsense conceptions of mental states contained in our present FOLK PSYCHOLOGY are, from a scientific point of view, radically defective. But we can expect that eventually a more sophisticated theoretical account will replace those folk-psychological conceptions, showing that mental phenomena, as described by current folk psychology, don't exist. Since that account would be integrated into the rest of science, we'd have a thoroughgoing materialist treatment of all phenomena. So this version of eliminativist materialism, unlike Rorty's, does not rely on assuming that the mental is non-physical.

But even if current folk psychology is mistaken, that doesn't show that mental phenomena don't exist, but only that they aren't the way folk psychology describes them as being. We could conclude they don't exist only if the folk-psychological claims that turn out to be mistaken actually define what it is for a phenomenon to be mental. Otherwise, the new theory would still be about mental phenomena, and indeed would help show that they're identical with physical phenomena. Churchland's argument, like Rorty's, depends on a special way of defining the mental, which we needn't adopt. It's likely that any argument for eliminative materialism will require some such definition, without which the argument would instead support the identity theory.

NECESSARY IDENTITY

Early identity theorists insisted that the identity between mental and bodily events was contingent, meaning simply that the relevant identity statements were not conceptual truths. That leaves open the question of whether such identities would be necessarily true on other construals of necessity.

Saul A. Kripke (1980) has argued that such identities would have to be necessarily true if they were true at all. Some terms refer to things contingently, in that those terms would have referred to different things had circumstances been relevantly different. Kripke's example is 'The first Postmaster General of the U.S.', which, in a different situation, would have referred to somebody other than Benjamin Franklin. Kripke calls these terms non-rigid designators. Other terms refer to things necessarily, since no circumstances are possible in which they would refer to anything else; these terms are rigid designators.

If the terms '*a*' and '*b*' refer to the same thing and both determine that thing necessarily, the identity statement '$a = b$' is necessarily true. Kripke maintains that the term 'pain' and the terms for the various brain states all determine the states they refer to necessarily; no circumstances are possible in which these terms would refer to different things. So if pain were identical with some particular brain state, it would be necessarily identical with that state. But Kripke argues that pain can't be necessarily identical with any brain state, since the tie between PAINS and brain states plainly seems contingent. He concludes that they cannot be identical at all.

This argument applies equally to the identity of types and tokens. Whenever the term 'pain' refers to a state, it refers to that state rigidly; similarly with the various terms for brain states. So if an individual occurrence of pain were identical with an individual brain state, it would be necessarily identical with it. Since they can't be necessarily identical, they can't be identical at all.

Kripke notes that our intuitions about whether an identity is contingent can mislead us. Heat is necessarily identical with mean molecular kinetic energy; no cir-

cumstances are possible in which they aren't identical. Still, it may at first sight appear that heat could have been identical with some other phenomenon. But it appears this way, Kripke argues, only because we pick out heat by our sensation of heat, which bears only a contingent tie to mean molecular kinetic energy. It's the sensation of heat that actually seems to be connected contingently with mean molecular kinetic energy, not the physical heat itself.

Kripke insists, however, that such reasoning cannot disarm our intuitive sense that pain is connected only contingently with brain states. That's because for a state to be pain is necessarily for it to be felt as pain. Unlike heat, in the case of pain there's no difference between the state itself and how that state is felt, and intuitions about the one are perforce intuitions about the other.

Kripke's assumption about the term 'pain' is open to question. As Lewis notes, one need not hold that 'pain' determines the same state in all possible situations; indeed, the causal theory explicitly allows that it may not. And if it doesn't, it may be that pains and brain states are contingently identical. But there's also a problem about a substantive assumption Kripke makes about the nature of pains, namely, that pains are necessarily felt as pains. First impressions notwithstanding, there is reason to think not. There are times when we are not aware of our pains, for example when we're suitably distracted. So the relationship between pains and our being aware of them may be contingent after all, just as the relationship between physical heat and our sensation of heat is. And that would disarm the intuition that pain is connected only contingently with brain states.

SUBJECTIVE FEATURES

Kripke's argument focuses on pains and other sensations, which, because they have qualitative properties, are frequently held to cause the greatest problems for the identity theory. Thomas Nagel (1974) traces the general difficulty for the identity theory to the CONSCIOUSNESS of mental states. A mental state's being conscious, he urges, means that there's something it's like to be in that state. And to understand that, we must adopt the point of view of the kind of creature that's in the state. But an account of something is objective, he insists, only insofar as it's independent of any particular type of point of view. Since consciousness is inextricably tied to points of view, no objective account of it is possible. And that means conscious states cannot be identical with bodily states.

The viewpoint of a creature is central to what that creature's conscious states are like because different kinds of creatures have conscious states with different kinds of qualitative property. But the qualitative properties of a creature's conscious states depend, in an objective way, on that creature's perceptual apparatus. We can't always predict what another creature's conscious states are like, just as we can't always extrapolate from microscopic to macroscopic properties, at least without having a suitable theory that covers those properties. But what a creature's conscious states are like depends in an objective way on its bodily endowment, which is itself objective. So these considerations give us no reason to think that what those conscious states are like is not also an objective matter.

If a sensation isn't conscious, there's nothing it's like to have it. So Nagel's idea that what it's like to have sensations is central to their nature suggests that sensations cannot occur without being conscious. And that in turn seems to threaten their objectivity. If sensations must be conscious, perhaps they have no nature independently of how we're aware of them, and thus no objective nature. Indeed, it's only conscious sensations that seem to cause problems for the identity theory (*see* SUBJECTIVITY).

The assumption that mental states are invariably conscious, like the supposition that they're non-physical, is basic to the Cartesian view. But sensations do occur that aren't conscious. A mental state's being conscious consists in one's being conscious of it in a way that's intuitively direct

and unmediated. But as already noted, distractions often make us wholly unaware of our sensations. Sensations that aren't conscious also occur in both subliminal perception and peripheral vision, as well as in more esoteric contexts. (*See* Weiskrantz, 1986.)

Sensations can, moreover, have qualitative properties without being conscious. Qualitative properties are sometimes called QUALIA, with the implication that we must be conscious of them; but we needn't be bound by that term's implications. Qualitative properties are simply those properties by means of which we distinguish among the various kinds of sensations when they're conscious. But a sensation's being conscious makes no difference to what its distinguishing properties are; its being conscious consists simply in one's being conscious of those properties in a suitable way. When a sensation isn't conscious its distinguishing properties seem to cause no difficulty for the identity theory. And since those properties are the same whether or not the sensation is conscious, there's nothing about those properties that undermines the identity theory. We would assume otherwise only if we held, with Nagel and Kripke, that sensations must all be conscious. (*See* Rosenthal, 1986.)

Perhaps multiple realizability refutes the type identity theory; but there are ample arguments that support the token identity theory. Moreover, the arguments against the token theory seem all to rely on unfounded Cartesian assumptions about the nature of mental states. The doctrine that the mental is in some way non-physical is straightforwardly question begging, and it's simply not the case that all sensory states are conscious. It is likely, therefore, that the identity theory, at least in the token version, is correct.

See also An Essay in Mind section 3.6.2.

BIBLIOGRAPHY

Armstrong, D.M. 1968. *A Materialist Theory of the Mind*. New York: Humanities Press.

Churchland, P.M. 1981. Eliminative materialism and the propositional attitudes. *Journal of Philosophy*, LXXVIII: 2, 67–90.

Davidson, D. 1970. Mental events. In *Experience and Theory*, ed. L. Foster and J. W. Swanson: Amherst, MA.: University of Massachusetts Press.

Kim, J. 1984. Epiphenomenal and supervenient causation. In *Midwest Studies in Philosophy, Vol 9*: Minneapolis: University of Minnesota Press.

Kripke, S.A. 1980. *Naming and Necessity*. Cambridge, MA.: Harvard University Press. Originally published in *Semantics of Natural Language*, ed. D. Davidson and G. Harman. Dordrecht: D. Reidel Publishing Company, 1972.

Lewis, D. 1980. Mad pain and Martian pain. In *Readings in Philosophy of Psychology, Vol. I*, ed. N. Block. Cambridge, MA.: Harvard University Press.

Lewis, D. 1972. Psychophysical and theoretical identification. *Australasian Journal of Philosophy*, L: 3, 247–58.

Nagel, T. 1974. What is it like to be a bat? *The Philosophical Review*, LXXXIII: 4, 435–50.

Putnam, H. 1975. The nature of mental states. *Mind Language and Reality: Philosophical Papers, Vol. 2*. Cambridge University Press.

Rorty, R. 1979. *Philosophy and the Mirror of Nature*. Princeton University Press.

Rosenthal, D.M., ed. 1991. *The Nature of Mind*. New York: Oxford University Press.

Rosenthal, D.M. 1986. Two concepts of consciousness. *Philosophical Studies*, XLIX: 3, 329–59.

Smart, J.J.C. 1962. Sensations and brain processes. In *The Philosophy of Mind*, ed. V. C. Chappell. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Weiskrantz, L. 1986. *Blindsight*. Oxford University Press.

DAVID M. ROSENTHAL

imagery Imagery has played an enormously important role in philosophical conceptions of the mind. The most popular view of images prior to this century has been what we might call 'the picture theory'. According to this view, held by such diverse philosophers as Aristotle,

Blackwell
Companions to
Philosophy

A Companion to the Philosophy of Mind

Edited by

SAMUEL GUTTENPLAN