Davidson on Mental Causation

Donald Davidson in "Actions, Reasons, and Causes" has given a closely reasoned explanation and defense of the commonsense theory of mental causation (Davidson 1968: 199 - 211). However, in another widely discussed article, "Mental Events," Davidson expounded a view of the nature of mental events – anomalous monism – that many critics claim has committed him to epiphenomenalism (Davidson 1979: 217 - 238). Davidson denies the charge, but he has been on the defensive ever since the appearance of "Mental Events." I argue that his critics are right if anomalous monism (AM) is regarded as an ontological claim, but that AM is no threat if it is regarded as a thesis about the practice of intentional explanation.

In "Actions, Reasons, and Causes," Davidson begins by stating his objective. He is concerned with "the relation between a reason and an action when the reason explains the action by giving the agent's reason for what he did." His term for this type of explanation is "rationalization." He says that he wants to defend the "common-sense position that rationalization is a species of ordinary causal explanation" (Davidson 1968: 199).

> A rationalization gives the primary reason for an agent's action. Whenever someone does something for a reason, . . . he can be characterized as (a) having some sort of pro attitude toward actions of a certain kind, (b) believing (or knowing, perceiving, noticing, remembering) that the action is of that kind. (Davidson 1968: 199)

Davidson then states and defends two theses about primary reasons. The first is:

> C1. R is a primary reason why an agent performed the action A under the description d only if R consists of a pro attitude of the agent toward actions with a certain property, and a belief of the agent that A, under the description d has that property. (Davidson 1968: 201)

Davidson's example is that of a person who flips a switch in order to turn on a light. The primary reason is the desire to turn on the light and

the agent's belief that the light can be turned on by flipping the switch. (I suspect that being an action entails that the action is caused by a primary reason.) So Davidson is prepared to conclude:

> C2. A primary reason for an action is its cause. (Davidson 1968: 205)

In the rest of the paper Davidson defends C2 against the objections of critics. (1) He shows that beliefs and wants, even though they are not events, but states or dispositions, are indispensable properties of anything that can be a cause of an action. (2) He shows that the "logical connection argument" has no force against the thesis that rationalizations are casual explanations. (3) He argues that first-person authority with respect to one's reasons does not disqualify them from being causes. (4) Hart and Honore had argued that "laws are involved essentially in ordinary causal explanations, but not in rationalizations" (Davidson 1968: 208). Thus rationalizations are not causal explanations (in the ordinary sense). Davidson's defense against this last objection foreshadows things that he will say in "Mental Events" that will throw doubt on whether rationalizations are ordinary causal explanations, as Davidson thinks they are.

Hart and Honore claim "The statement that one person did something because, for example, another threatened him, carries no implication or covert assertion that if the circumstances were repeated the same action would follow" (Davidson 1968: 208). In other words, there is nothing like a causal law here. Some writers have suggested that there are rough generalizations connecting reasons and actions, and that (theoretically) these could be improved. Strangely, Davidson does not accept this kind of defense. He says that the suggestion is "delusive" and that rationalizations "cannot be sharpened into the kind of law on the basis of which predictions can reliably be made" (Davidson 1968: 208). Objections arise. Is the impossibility a theoretical impossibility? Is the impossibility true of rationalizations, but not rue of other causal generalizations? If so, then it may not be true that rationalizations are ordinary causal explanations.

But other things Davidson says strengthens the claim that rationalizations are ordinary causal explanations. Suppose a window breaks because it was struck by a rock. We are very certain that it was the impact of the rock that broke the window, but can anyone point to a strict law that governs the cause? In general, we are much more certain of singular causal statements than we are of causal generalizations. Reflection along these lines tends to diminish the difference between everyday causal explanations and rationalizations. Does this mean that Hume was wrong when he said that

singular causal statements entail laws? Davidson says that Hume's claim is ambiguous.

> It may mean that "A caused B" entails some particular law involving the predicates "A" and "B," or it may mean that "A" caused "B" entails that there exists a causal law instantiated by some true descriptions of A and B. (Davidson 1968: 209)

Davidson accepts the second, weaker version of Hume's claim which does not require that a singular causal statement be defended by defending any causal law. He recognizes that it is the weaker version of Hume's doctrine that fits most singular causal statements, not just rationalizations. Still, I suspect that a wedge has been made that is partly responsible for the doctrine of anomalous monism that first appears in "Mental Events."

In "Mental Events" Davidson seems to be more concerned with physical causes of mental events than he is with mental causes of physical events. He gives more attention to perception and the threat to freedom posed by the determination of behavior by physical causes than he does to the mental causation of behavior. Nevertheless, subsequent discussion has settled almost entirely on the latter. The doctrine of anomalous monism (AM) presented in the article has seemed to some to commit Davidson to epiphenomenalism.

Davidson notes that there is an apparent contradiction between the facts that mental events and states play a causal role in the physical world and the absence of causal laws connecting the mental and the physical (Davidson 1979: 218). He says that the apparent contradiction stems from three principles. The first principle asserts that mental events interact causally with physical events. Mental events cause physical events, and physical events cause mental events. The second principle – the principle of the nomological character of causality – requires that events related as cause and effect fall under strict laws, and the third principle – the principle of the anomalism of the mental – is that there are no strict laws connecting mental and physical events (Davidson 1979: 219).

Davidson combines the anomalism of the mental with a version of the identity theory to obtain AM. His identity theory seems to be a strict token-identity theory. I call it strict, because he seems to be denying the very possibility that there could be type-identity between the mental and the physical. He cites three writers – Charles Taylor, Jaegwon Kim, and J. J. C.

Smart – who he thinks are representative of identity theorists. What they all falsely believe is that the identity of a mental with a physical event entails that there is a correlating law connecting that type of mental event with that type of physical event. Davidson denies this and says that it is possible to know that a mental event is identical with a physical event without knowing which physical event it is identical with (Davidson 1979: 223).

Davidson ends his article by claiming that AM shows how it is possible for thought and purpose to be causally efficacious and at the same time free from the operation of law (Davidson 1979: 235 - 236).

Although Davidson denies that there are psycho-physical laws, he says that mental characteristics may supervene on physical characteristics, and he takes such supervenience to mean that "there cannot be two events alike in all physical respects but differing in some mental respect" (Davidson 1979: 225).

Davidson's explanation of how AM reconciles his three original principles is puzzling. He says that cause and effect is a relation between two events no matter how described. But laws "are linguistic; and so events can instantiate laws, and hence be explained and predicted in the light of laws, only as those events are described in one way or another way." I agree that how events are described has nothing to do with whether one is the cause of the other. The truth of the causal statement will depend on which events are referred to. However, laws of nature are not linguistic facts, either. They are as independent of linguistics as causes are. Evidently, Davidson does not make a distinction between law-statements and laws, but he should.

In "Thinking Causes," Davidson attempts to clarify what AM is, tries to show how the three premisses from which he argues to AM are consistent with one another, and defends AM against the charge that AM makes the mental epiphenomenal (Davidson 1993: 03 - 17). First, consider his clarification (and perhaps modification) of what he had said in "Mental Events." In the opening paragraph he provides another definition of AM.

> AM holds that mental entities . . .are physical entities, but that mental concepts are not reducible by definition or natural law to physical concepts. The position is in a general way familiar: it endorses ontological reduction but eschews conceptual reduction. (Davidson 1993: 03)

He may endorse ontological reduction, but he does not provide even the faintest suggestion of how it is possible. It is perhaps this failing that causes some of his critics to treat Davidson as if he were an opponent of ontological reduction.

He says that he will defend not only AM, but AM in conjunction with the three premisses from which it is derived and the doctrine of supervenience (AM + P + S). "Supervenience" is then defined as follows: "a predicate p is supervenient on a set of properties S if and only if p does not distinguish any entities that cannot be distinguished by S" (Davidson 1993: 04). The definition implies (Davidson says) that the extension of the supervenient predicate is determined by the extension of the subvenient predicates(s). He denies, however, that the definition implies that p can be nomologically reduced to S.

Davidson's primary concern is the defense of AM + P + S against the charge that it implies mental property epiphenomenalism (MP-epiphenomenalism). In the rest of the paper I will examine this criticism and Davidson's response to it. Although I will conclude that the criticism is justified under a reasonable interpretation of AM, I will suggest that the criticism can be avoided if AM is regarded as a thesis, not about ontology, but as a thesis about psychological explanation.

Kim is one of the writers who have claimed that there are epiphenomenalist tendencies in AM. According to Kim: "The fact is that under Davidson's anomalous monism, mentality does no causal work. Remember: on anomalous monism, events are causes only as they instantiate physical laws, and this means that an event's mental properties make no causal difference" (Davidson 1993: 05).

Davidson tries to rebut Kim in various ways, but perhaps it can be made fairly clear why AM has seemed to some writers to lead to MP-epiphenomenalism. Suppose some definite description d refers to event X, that X is mental under d, that some other definite description d′ refers to event Y, and that X is the cause of Y. Then, necessarily, X instantiates some property p and Y instantiates some property p′, and any instantiation of p will cause an instantiation of p′. But there are no strict causal laws relating mental properties to anything else. So p could not be the mental property of X. Like all mental properties, the mental property of X is causally irrelevant.

Davidson seems at times to misunderstand how AM is being criticized. He may think that his critics are claiming that mental events cannot cause physical events, but what they are claiming is that it is not by virtue of their mental properties that mental events cause physical events. Davidson has always emphasized that the truth of a causal statement depends on which events are described, not how they are described; *i.e.*, he asserts the independent reality of causes and effects. However, most of his critics are realists with respect to causes and effects, and when they say that he has neglected the role that properties play in causation, they are not even

talking about descriptions. Perhaps the following will suggest the confusion that Davidson may labor under.

> . . .if causal relations and causal powers inhere in particular events and objects then the way those events and objects are described, and the properties we happen to employ. . . to characterize them, cannot affect what they cause. Naming the American invasion of Panama "Operation Just Cause" does not alter the consequences of the event. (Davidson 1993: 08)

That the way we describe an object and the properties we cite to characterize it are causally irrelevant does not mean that the properties it *has* are causally irrelevant.

Davidson is concerned with showing that the supervenience of mental properties on physical properties is non-reductive. Perhaps he thinks that reduction of the mental to the physical would imply the falsity of AM. However, it seems to me that reduction of the mental to the physical would be one way of avoiding MP-epiphenomenalism. But I need to examine some of the ways that supervenience has been used.

Historically the term was first used in connection with value qualities. R. M. Hare was the first to use the term, but Davidson gives G. E. Moore credit for originating the concept. According to Davidson, Moore asserts that "something is good only because it has properties that can be specified in descriptive terms, but goodness can't be reduced to a descriptive property" (Davidson 1993: 04). Here are some examples of properties supervening on other properties. A woman is described as dumpy because she is four feet seven inches in height and three feet nine inches in width. A man is called good because he is honest, kind, and useful to society. A knife is called good because it is sharp and durable. One writer says that being morally outrageous is a property that supervenes on Nixon's behavior during Watergate.

One thing to be noticed about all these examples is that the supervenience is conceptual in nature – we can see by reflection that the supervenient property supervenes on the subvenient properties. So Moore (and Davidson) must be wrong when they deny that conceptual reduction of value qualities is possible. Of course, the relation between brain events or states and their mental counterparts is nothing like this. It could not be, because none of us, including the experts, have the faintest idea of which brain entities correlate with which mental entities. We need to find examples

of supervenient properties that may be more similar to the way that mental properties are related to neural properties.

Consider the relation between a disposition and its basis. According to the reductivist theory, the disposition and its basis are identical. Suppose some object has a tendency to shatter when struck with sufficient force and it has this tendency by virtue of its crystalline structure; then the descriptions "tendency to shatter" and "crystalline structure" refer to the same condition in the object. "Tendency to shatter" and "crystalline structure" differ in meaning, but when they are used to refer to some condition of the object they refer to the same condition. Notice, however, that in one important respect the relation between the disposition and its basis is unlike the relation between a value quality and the factual matters on which it depends. The latter, but not the former, is conceptual. The concept of the disposition does not include a concept of what its basis is.

Now consider an over simplified statement of the identity theory. According to the theory, any mental state or event is identical to some condition of the brain. The relation is nearly parallel to the relation between a disposition and its basis. Like dispositions, mental concepts do not include concepts of what their bases are, so mental concepts cannot be reduced to neural concepts. It is a conceptual requirement that a disposition has a basis, but I do not know whether there is a similar requirement for mental entities. However, I am certain that this possibility is not conceptually excluded.

Suppose the identity theory is correct. Then I think it could be said that in some sense the mental is reducible to the physical. Perhaps it would be wrong to say that this is a reduction of mental properties to physical properties. What we think of as properties is too closely tied up with the meanings of words. A tendency to shatter is not the same property as crystalline structure because "tendency to shatter" does not mean what "crystalline structure" means. However, it still seems obvious that the two terms refer to the same reality. And the same could be said for mental events if the identity theory is correct. "Feeling a pain" does not have the same meaning as "the firing of neuron X." So it could not be said that the mental property had been reduced to the physical property. Nevertheless, the two terms would be referring to the same reality, for there would be nothing to the feeling of pain over and above the firing of neuron X.

An identity theory has several advantages of which I will mention only two. First, there would be no threat to the closed causal character of the physical world, because mental events would be physical events and every aspect of the mental would at the same time be an aspect of the physical. More importantly, the charge of MP-epiphenomenalism directed against Davidson would carry no weight. The mental would be so intimately bound

up with the physical that it would be impossible for epiphenomenalism to find a toehold.

This type of reductionism should not be confused with eliminative material which holds that psychological explanations of behavior should be replaced by neurological explanations that make no mention of intentional states. Nothing is implied about how explanations should proceed. And it is entirely consistent with the identity theory that there should be causal laws in which X causes Y in virtue of the mental properties of X. More specifically, intentional explanations are consistent with the theory. Some people (including Davidson) may have thought that a reductive identity theory implies eliminative materialism, but they would be wrong.

Lynne Baker proposes:

> . . .to perform a methodological about-face. Instead of beginning with a full-blown metaphysical picture we should begin with a range of good explanations . . . In the spirit of G. E. Moore, I think our grounds for the claims that reasons sometimes explain behavior are much stronger than any grounds for a metaphysical premiss that would lead to a contrary conclusion. (Baker 1993: 95)

Davidson's "Actions, Reasons, and Causes" is an account of common sense explanations of behavior, but in "Mental Events" he is more concerned with metaphysics. One of his metaphysical assumptions is the assumption that if one event C is the cause of another event E then there is some description of C and some description of E that together instantiate a strict law on those descriptions. This assumption together with the further assumption that there are no strict laws relating mental properties to anything else is the basis of the charge that AM commits Davidson to MP-epiphenomenalism. But what makes him think that C can be the cause of E only if a strict law is instantiated on some descriptions of C and E? There is a regulatory principle of both science and common sense that searches should be conducted for exceptionless laws, but it is only an assumption that there are strict laws that apply to every kind of event. At any rate, strict laws do not play a prominent role in either everyday explanation or the special sciences. So Davidson should let this metaphysical assumption go. It only gets him in trouble.

If metaphysics is put aside and emphasis is placed on Davidson's account of rationalization, "Mental Events" can be viewed as a valuable supplement to the earlier article. Some of the things he says are true and relevant. For instance, it is true that conceptual reduction of the mental to

the physical is impossible. And, although he has not shown that ontological reduction is theoretically impossible, it is important that nothing remotely resembling ontological reduction is going on now. But assume that in some remote future ontological reduction has made considerable headway. Even then, I suspect that explanations of behavior will go on in about the way they do now.

References

Baker, Lynne Rudder. 1993. Metaphysics and Mental Causation. In *Mental Causation*. J. Heyl and A. Mele, eds. Oxford: Oxford University Press.

Davidson, Donald. 1968. Actions, Reasons, and Causes. In *An Introduction to Philosophical Enquiry*, J. Margolis, ed. New York: Alfred A. Knopf.

------. 1979. Mental Events. In *Philosophy As It Is*. T. Honderich and M. Burnyeat, eds. New York: Penguin Books.

------. 1993. Thinking Causes. In *Mental Causation*, J. Heyl and A. Mele, eds. Oxford: Oxford University Press.