Elements of Mind

An Introduction to the Philosophy of Mind

Tim Crane
is non-spatial. It would be an interesting historical task to determine which kinds of dualism advocated by the philosophers of the past fall into which category, but there is no room for this task here. My point is that if we can make sense of the idea of a substance dualist view which does allow for the existence of the mind in space, then Fodor’s first point is not relevant to a general attack on substance dualism.

However, it turns out that we would not need this first point against dualism if Fodor’s second point is sound: that causation by non-physical substances would violate the laws of physics. For if this second point is true, then whatever non-physical substances turn out to be, they cannot have effects in the physical world unless the laws of physics are false. And in fact, it turns out that this criticism is even more powerful than this: for, as we shall see, it is an objection not just to substance dualism, but to any form of dualism whatsoever. So if the second point is sound, we do not need the first. It is important to emphasize this, since it is sometimes assumed that the arguments against dualist causation are essentially of the first kind. If I am right, this is a mistake.

But what exactly is Fodor’s second point? The objection is not that the dualist’s conception of interaction between the mental and the physical is unintelligible, that no sense whatsoever can be made of it. The claim is rather that the apparent fact of mental causation is inconsistent with other things we know. So let us start with the idea that the mind affects the body, and ask: with what known fact about the physical world is this incompatible? This requires a brief digression into physics, the physical, and the influential doctrine known as physicalism.

12. Physics and physicalism

Dualism, as its name suggests, classifies things in the world by number: it says there are two things, or two kinds of thing. Therefore it is naturally contrasted with monism, which says there is one thing, or one kind of thing. Monisms have been traditionally divided into two kinds: idealism, which says that all is mental, and materialism, which says that all is material.

But contemporary philosophers talk about a contrast and opposition between dualism and physicalism. Is there a difference? Sometimes the word ‘physicalism’ is used as a synonym for ‘materialism’, and ‘physical’ used as a synonym for ‘material’. This is perfectly natural insofar as physics is the science of matter, but it is nonetheless possible to make some useful distinctions between monism, materialism, and physicalism. Making these distinctions will help us to understand the objection to a dualist account of mental causation.
Monism is a commitment to what we might call one-ness: the world is one, of one nature. A truly monistic view is Spinoza’s monism, which holds that there is only one substance, which can be called God or Nature. As we saw above, Spinoza adopted the traditional conception of substance as that which can exist in and of itself, and is not dependent on anything else. So it is not hard to see how he came to the conclusion that there can be only one substance, since God is the only thing which is genuinely independent of everything else. Contemporary versions of monism tend to say not that the world itself is one substance, but that everything in the world is made of the same sort of thing: materialism says that the world is made of matter, idealism says that it made of ideas, and Russell’s ‘neutral monism’ of the 1920s says that the world is of one nature, while remaining neutral on whether it is mental or material.

Here I am interested in making a distinction between materialist monism and physicalism. A materialist monist holds that everything is material, that is, made of matter. But it is obvious that physics (the science of matter, among other things) says that there are many things in the world which are not made of matter: there are forces, waves, fields, and so on. A physicalist, traditionally, is someone who gives a certain kind of authoritative role to physics. This role is partly epistemological—physics has an authority in telling us what to believe—and partly ontological—physics has an authority in telling us what there is. This conception of the special role of physics has complex origins. But at its heart is the idea that physics is the science which aims at what W. V. Quine calls ‘full coverage’. Physics aims, by using exact quantitative methods, and its categories of mass, energy, force, and so on, to give an account of the properties and behaviour of everything which has a spatial and temporal position. The laws of physics are laws which are intended to be true of all objects in space and time. There are no objects which are exempt from these laws. Let’s call this claim, ‘the generality of physics’:

\[
\text{The generality of physics: All objects and events in space-time have physical properties, and the laws of physics govern or describe the behaviour of all objects and events in space-time.}
\]

The generality of physics is a monist principle. But it is not yet physicalism. For physicalism is the view that the physical story told by these laws is, in a sense, the whole story.

And full coverage is one thing, the whole story is another. Physicalism is the view that physics is the whole story, but this does not follow from the fact that physics has full coverage. (Thus it could be misleading to call a theory which unifies the fundamental physical forces a ‘theory of everything’ just because everything is subject to these forces.) In what sense, though, is physics supposed to tell the whole story, if ‘whole story’ does not just mean ‘full
coverage? Physicalism says that physics tells the whole story about the causation of physical events: that is, events which have physical properties or features. According to physicalism, everything physical which happens, everything which is an effect, must be a result of purely physical causes in accordance with physical law. This doctrine is a doctrine about causation. Following David Papineau, I shall call it ‘the completeness of physics’:

The completeness of physics: Every physical event has a physical cause which is enough to bring it about, given the laws of physics.

The label ‘completeness’ should not lead us to think that physics is a complete science in the sense that it has been completed—that physics is ‘finished’. (In any case, it is not clear what it would be for a science to be ‘finished’; the idea reminds me of a cartoon I once saw, with a picture of a road leading to a city, and a sign on the road saying something like: ‘No more building; this city is finished.’) Rather, the idea behind the phrase is that physical causes are complete in the sense of being enough to bring about all physical effects.

The completeness of physics should also be distinguished from the claim that physics can explain everything. Consider the view that David Lewis calls ‘the explanatory adequacy of physics’:

there is some unified body of scientific theories of the sort we now accept, which together provide a true and exhaustive account of all physical phenomena. They are unified in that they are cumulative: the theory governing any physical phenomenon is explained by theories governing phenomena out of which that phenomenon is composed and by the way it is composed out of them. The same is true of the latter phenomena, and so on down to fundamental particles or fields governed by a few simple laws, more or less as conceived in present-day theoretical physics.

This is a stronger claim than the completeness of physics, because one could hold the completeness of physics and deny the explanatory adequacy of physics, but not vice versa. We may reject the idea of the explanatory adequacy of physics because we may believe (as seems plausible) that the different sciences have their own different explanatory domains, which they treat in their own way, in terms of their own concepts and principles. Explanations in biology, for instance, appeal to biological concepts and categories, and no one believes that they can really be expressed in the language of fundamental physics. Or that they need to be. Yet this explanatory autonomy for biology is compatible with both the generality of physics—biological interactions are among things with physical properties, and no biological interaction is in conflict with the laws of physics—and the completeness of physics, since the completeness of physics says that whatever physical events occur have a physical causal history which fixes their occurrence.
The completeness of physics may be illustrated with a theological image. Imagine God creating the universe. God has to decide how to set things up, to arrange the matter at the beginning of the universe and to choose the laws according to which the matter will behave and so on—in such a way that he gets the universe he wants. The completeness of physics says that in order to get every physical effect in the universe, God does not have to do anything else except set up the physical laws and initial conditions. All God has to do to make anything physical happen is to set the physical part of the world in motion.

The image seems to imply determinism: the doctrine that the present is completely fixed by the nature of the past and its exceptionless laws, in such a way that if the universe were to be started all over again, with the same initial conditions and the same set of (deterministic) laws, it would have the same particular history. But in fact, the completeness of physics does not imply determinism, and it is consistent with an indeterministic view of the physical universe, according to which physical events happen only with a certain probability. According to indeterminism, it does not follow that if the universe were started all over again with the same laws of nature, it would have the same particular history. There is only a certain chance that it has the same particular history, where chance is understood as objective physical probability. Whether the universe is deterministic, or (as contemporary physics seems to suggest) indeterministic is a question which is independent of the truth or falsehood of the completeness of physics. Strictly speaking, then, when discussing the completeness of physics, we should express it in the way Papineau does, when he says:

*The completeness of physics* 2: Every physical event is determined, or has its chance determined, by purely physical causes in accordance with physical law.

But no harm will be done if we simplify and talk as if determinism were true in this chapter.

Physicalism has been defined in many ways. What it (or the completeness of physics) really amounts to depends almost entirely on what ‘physical’ means. I have been taking the meaning of ‘physical’ to be given by the content of physical science, where physical science is the science which aims at full coverage. The nature of this science is not something which can be established by purely *a priori* reflection; it is an empirical question what the content and scope of physics actually is. This marks one difference between physicalism and older forms of materialism, which fixed the content of its doctrine in a relatively *a priori* way: for instance, by saying that everything is material, and matter is solid, impenetrable, conserved, interacting deterministically and only on contact. Since modern physics has shown this conception of matter to be wrong in perhaps every respect, it is reasonable for a materialist to become
a physicalist, and take the approach: ‘rather than say a priori what the world of matter must be like, I will rather let physics, the science of matter, tell us what matter, and the rest of the world, is like.’ The point of calling yourself a physicalist rather than a materialist is chiefly to express this attitude to physical science.

This does mean that there is an open-ended character to the doctrine of physicalism itself. Should we say that the content of physicalism is fixed by present-day physics, or by some ideal future physics? Either gives problems: for present physics is incomplete and may be false in certain respects; yet who knows what is going to be the content of the ideal future physics? It looks as if physicalism is either obviously false (if physics is today’s physics) or empty (since who knows what will be in the physics of the future).

This is a nice problem, but physicalism can solve it. Physicalism asks us to address the ontological question in this way: see what physics says there is, and then commit yourself to that kind of thing being all there is. As time develops, it may be that your commitments develop too. But this is just a reflection of the fact that you have no standard (other than physics) from which to answer the question of what there is.

This open-ended character does, however, limit what physicalists should permit themselves to say. They should not allow themselves to say, for example, that a physicalist must hold that there are no ghosts. Certainly current physics holds that there are no ghosts, but if it were discovered that irreducible ghosts were needed to explain certain physical phenomena, then ghosts would be physical by this definition. Or consider parapsychology. If it turned out that there is good evidence that parapsychological phenomena—telekinesis, telepathy, and so on—were needed in order to explain certain physical effects, then these phenomena would have to become part of the realm of the physical. These possibilities are, of course, unlikely—not least because there is no uncontroversial, undisputed, solid evidence for parapsychological phenomena—but that does not affect the point of principle, which is that if we give ‘physical’ the meaning it has in ‘physical science’ then we cannot say a priori what the physical is, since we cannot say a priori what physical science is. But, I claim, physicalists can reasonably ignore the remote possibilities just discussed.

Someone could understand the physical in other terms. But none of these alternative understandings is satisfactory, since none can make any sense of the current disputes; in particular, the dispute we are considering about mental causation. For example, someone might say that the physical is what exists in space and time—but how would they then rule out the dualism considered above (§3) as being physicalism? Or they might say that the physical is the causal—but given the evident fact of mental causation, this makes mental phenomena physical by definition. Such definitions of
'physical' make physicalism virtually trivial; and while some may be happy with this, it certainly would not satisfy physicalists, who think of their doctrine as substantial and informative.

So what can physicalism say? If it is not an empty doctrine, devoid of all genuine content, what metaphysical claim can physicalism allow itself to make? This is where the completeness of physics comes in. Physics, as it stands, and as it is likely to remain, attempts explanations of why things happen—things like the collision of particles, the motion of projectiles, and so on. Physicists claim that they can explain such events (often they say ‘explain in principle’) in terms of their dynamical equations and in terms of a small number of basic concepts—force, charge, momentum, acceleration, and so on. A metaphysical generalization of this is the completeness of physics. What would it be, then, to deny the completeness of physics? It would be to hold that some physical effects—some effects of the same general kind as the kind of which physics treats—would not come about were it not for the presence of other causes, other causes which are non-physical in the sense of not being the subject-matter of physical science. This is something no physicalist—no one who shares the view that physics has a unique ontological and epistemological authority—can believe. I therefore claim that the completeness of physics must be an essential component of any physicalist view, a necessary condition for any non-trivial form of physicalism.

What has this got to with Fodor’s second point? Fodor says that mental causation would violate the laws of the conservation of energy and mass. But it would be preferable to see Fodor’s point as an expression of the completeness of physics: the reason that energy would not be conserved in a mental-physical interaction is that every physical effect must come about through purely physical causes. Mental causation would therefore have to introduce ‘more energy’ into the physical world, thus violating the conservation laws. The world of physical effects must be causally closed, according to physicalism.

13. The problem of mental causation for dualists

The completeness of physics is a necessary condition for physicalism, but it is not a sufficient condition. For one could hold the completeness of physics and still believe that there are mental things and properties, but that they have no effects. This is epiphenomenalism: the doctrine that mental states and properties have no physical effects. But the dualist views we are considering deny epiphenomenalism: they are interactionist views. Given this, we can now state the problem of mental causation for dualism, which is one part of what is now known as the mind–body problem.
The problem comes from the conflict between the existence of mental causation and the completeness of physics. Assume:

(1) Mental phenomena have effects in the physical world;

add the completeness of physics:

(2) All physical effects have physical causes which are enough to bring those effects about;

and it is easy to see, in general outline, how the conflict arises. How can the mental cause bring a physical effect about if the physical cause is itself enough to bring it about?

It might seem as if one could answer by saying: the mental cause is just an extra cause. It is something added on to the physical cause. To see what might be wrong with this, we need to make explicit another assumption, which is the following:

(3) Mental and physical causes do not overdetermine their physical effects.

Causal overdetermination is when an effect has more than one cause, and each event would have caused the effect if the other one had not done so. Consider the assassination of a tyrant by two assassins. Each assassin shoots the tyrant, and so is a cause of the tyrant’s death. But the assassination is set up in such a way that the shootings are independent: either would have killed the tyrant if the other had failed. The death of the tyrant is overdetermined by the shootings.

(Overdetermination in this sense must be distinguished from an event’s having more than one cause. Consider the Second World War: one of its causes was Hitler’s invasion of Poland, another was the invasion of Czechoslovakia. Moving further back, another cause might be thought to be the resentment felt by Germany after the Treaty of Versailles. All of these are plausibly among the many causes of the war, but of none of them is it true that any of them would have brought about the war if the others had not been there.)

Does the possibility of overdetermination make sense? Some would say that it does not. Those who believe in the counterfactual analysis of causation are committed to the following claim:

(C) If A caused B, then if A had not been the case, B would not have been the case.

Applied to our example, we find that the counterfactual analysis entails that neither assassin’s shot can be a cause of the tyrant’s death! Because: of neither shooting is it true that if it hadn’t been done, the death would not have occurred. We could either conclude that, despite appearances, overdetermination
like that in the case of the tyrant’s death cannot really happen; or we could conclude that the counterfactual analysis © is false.

This is an important issue in metaphysics; but fortunately, we do not need to settle it for present purposes. For even if overdetermination is possible in these very unusual circumstances, and © is strictly speaking false, it is nonetheless very implausible to suppose that this is the way mental and physical causes relate to each other. For if they did, then every time a mental state had some effect in the physical world, the completeness of physics guarantees that there would be a cause in one’s brain which is itself enough to bring about that very same effect. It then looks like a coincidence that my body manages to co-ordinate so well, given these distinct causes of its motions.²⁸ Yet the idea that one’s control of one’s body is in this way coincidental is in conflict with everything we know about the causation of behaviour—either from our own experience, or from common sense. That is the justification for (3) above.

The problem of mental causation for dualism, then, is how to reconcile the existence of mental causation with the completeness of physics and the denial of the general overdetermination by mental and physical causes. This way of expressing the problem brings to light something very important which is sometimes overlooked: the problem does not arise because of something about the mental. It’s because of a fact about the physical world: the completeness of physics. It’s the assumed nature of the physical world which generates this aspect of the mind–body problem.

But didn’t we assume substance dualism in setting up this problem? Yes, we did, since this is the way the problem has traditionally been formulated. But it turns out that this assumption was not essential. For all we are really assuming about the mental is that mental phenomena have physical effects, and that these mental causes are distinct from the physical causes of the same effects. Suppose that we rejected substance dualism, and accepted only a dualism of properties. Then, so long as we are thinking of the properties of things as causes (more on this below), the problem will still arise as above. The source of the problem is the conflict between the completeness of physics and causation by mental causes which are distinct from physical causes.

Since this problem is the focus of our discussion for most of the rest of this chapter, we can put the question of substance dualism to one side. As far as mental causation goes, substance dualism is a red herring. So rejecting substance dualism for the (admittedly weaker) property dualism will not help in solving this problem. For it still treats mental and physical causes as distinct. What obviously would help is to deny this: that is, to deny that there are two causes here. This would be to accept the identity theory of mental and physical causes:

(4) Mental causes are identical with physical causes.
This solves the problem. Since there are not two causes, but one, there is no risk of overdetermination. Mental causes are exactly the same entities as certain physical causes—in the brain, no doubt. This is the *identity theory* of mind and brain (or mind and body), and for a while it was the dominant physicalist theory of mind. I believe that the identity theory certainly solves the problem we have been discussing. But is it an independently plausible theory?

### 14. The identity theory

The argument just presented for the claim that mental entities are identical with physical entities needs two clarifications. First, the argument as it stands only gives us a reason for identifying with brain states those mental entities which are causes of physical events. If there are mental entities which have no physical effects (mental epiphenomena) then there is, as things stand, no reason to identify them with physical entities. Although there are those who hold that there are mental epiphenomena, many physicalists will say that all mental phenomena do have physical effects in one way or another, so we know enough to rule out epiphenomenalism. One approach which would guarantee this is the *functionalist* approach to mental phenomena, which says that mental phenomena are individuated—distinguished from one another—by their causal roles. The idea here is that the concept of a mental state is a causal concept: the concept of perception, for instance, is the concept of a state of mind which has certain typical causes (perceptible objects and events in the environment, say) and certain typical effects (the perception that it is snowing, say, typically causes the belief that it is snowing). Functionalism therefore holds that it is in the nature of certain mental states to have certain effects; therefore there can be no mental epiphenomena. Combining this with the completeness of physics and the denial of overdetermination yields the identity theory—and this is, in fact, how some functionalists have argued.²⁹

The second clarification is that there are two ways in which causes can be conceived, and each way corresponds to a different kind of identity theory. Consider a simple causal interaction such as a brick breaking a window by being thrown at the window. Davidson has argued that causes and effects are events, where events are a kind of particular (see §10). According to Davidson, the event which is the cause of the window’s breaking is the brick’s hitting the window. This is a particular event which can be described in many ways: it could be described as ‘the throwing of the brick’ or as ‘the throwing of a brick made in Walthamstow’ or as ‘the throwing of a red brick’. Each of these descriptions can be a description of the same event, and therefore of the same cause.