1

The Rise of Physicalism

DAVID PAPINEAU

> No one could seriously, rationally suppose that the existence of antibiotics or electric lights or rockets to the moon disproves . . . mind-body dualism.
>
> Stephen R.L. Clark (1996)

Introduction

In this chapter I want to discuss the way in which physical science has come to claim a particular kind of hegemony over other subjects in the second half of this century. This claim to hegemony is generally known by the name of *physicalism*. In this chapter I shall try to understand why this doctrine has come to prominence in recent decades. By placing this doctrine in a historical context, we will be better able to appreciate its strengths and weaknesses.

As a preliminary, note that contemporary physicalism is an ontological rather than a methodological doctrine. It claims that everything[1] is physically constituted, not that everything should be studied by the methods used in physical science. This emphasis on ontology rather than methodology marks a striking contrast with the 'unity of science' doctrines prevalent among logical positivists in the first half of the century. The logical positivists were much exercised by the question of whether the different branches of science, from physics to psychology, should all use the same method of controlled observation and systematic generalization. They paid little or no attention to the question of whether everything is made of the same physical stuff.

By contrast, physicalism, as it is understood today, has no direct methodological implications. Some physicalists uphold the view that all sciences should use the "positivist" methods of observation and generalization. But as many would deny this. You can be a physicalist about biology, say, and yet deny that biology is concerned with laws, or a physicalist about sociology, and yet insist that sociology should use the method of empathetic verstehen rather than third-person observation.

This methodological liberalism goes with the fact that the ontological claims of fin-de-siècle physicalism are often carefully nuanced. If physicalism simply meant type-type physical reduction, of the kind classically characterized in Ernst Nagel's *The Structure of Science* (1961), then methodological unity of science would arguably follow, in principle at least, from physicalism. But physicalism today clothes itself in various subtler shades. We have physical supervenience, physical realization, token-token physical identity, and so on. These more sophisticated doctrines leave plenty of room for different sciences to be studied in different ways.

But I am already drifting away from the main subject of this chapter. My concern here is not to distinguish the different species of physicalism, though I shall touch on this in passing later, but to try to understand the reasons for physicalism of any kind. Why have so many analytic philosophers in the second half of the twentieth century suddenly become persuaded that everything is physical?

Fashions and Arguments

It certainly wasn't always so. Perhaps the easiest way to highlight the recent shift in thinking about physicalism is to recall a once-heated mid-century debate about the status of psychological explanation. In contemporary terms, this debate was about the scientificity of 'folk psychology.' On the one side were those, like Carl Hempel and A.J. Ayer, who argued that 'reasons are causes.' By this they meant that psychological explanations are underpinned by empirical generalizations, implicit in everyday thought, which link psychological states such as belief and desire to subsequent behavior. Opposed to Hempel and Ayer were thinkers such as William Dray, and Peter Winch, who argued that the links between reason and action are "logical" or "meaningful," not empirical (Hempel [1942]; Ayer [1969]; Dray [1957]; Winch [1958]).

In one respect this old debate is still up to date. It concerned the question of whether everyday psychological thinking is suitable for incorporation in a scientific psychology – whether folk psychology is a 'proto-science,' as it is sometimes put – and this question is still very much a live issue. But at another level the old debate is now quite outmoded. This is because the participants in the old debate showed little or no interest in the question of how the mind relates to the brain. They wanted to know whether there are testable, empirical laws linking mental states to behavior. But they seemed to see no connection between this issue and the question of the relation of mental states to brain states.

Nowadays, by contrast, everybody has a view on this latter question. Indeed nearly all analytic philosophers in this area, including those who side with Dray and Winch against the scientificity of commonsense psychology, now accept that the mind is in some way constitutively connected with the brain. (Thus consider Donald Davidson. He is the modern champion of the Dray-Winch view that the explanatory links between reason and action are a *sui generis* matter of rational understanding, not scientific law. Yet he made his name by arguing that, even so, 'reasons *are* causes.' In effect, his contribution was to show how the Dray-Winch methodological denial of psychological *laws* could be combined with a physicalist commitment to mind-brain *constitution* (Davidson [1963]).)

This transformation of the old 'reasons and causes' debate happened very quickly. Until the 1950s the issue was purely about lawlike patterns. The issue of mind-brain identity was not on the agenda. Then suddenly, in the 1950s and 1960s, a whole stream of philosophers came out in favor of physicalism. First there were Herbert Feigl and the Australian central state materialists, and they were followed in short order by Donald Davidson, David Lewis, and functional state theorists such as Hilary Putnam. While the old 'reasons and causes' issue continued to be debated, from now on this debate took place within the larger context of physicalist assumptions about the mind-brain relation (Feigl [1958]; Place [1956]; Smart [1959]; Armstrong [1968]; Davidson [1963], [1970]; Lewis [1966]; Putnam [1960]).

Why exactly did physicalism come to prominence in this way in the 1950s and 1960s? Those antipathetic to physicalism sometimes like to suggest that the emergence of physicalism is essentially a matter of fashion. On this view, the rise of physicalism testifies to nothing except the increasing prestige of physical science in the modern weltanschauung. We have become dazzled by the gleaming status of the physical sciences, so the thought goes, and so foolishly try to make our philosophy in its image. (Thus Stephen Clark says, in the sentence immediately following the quote at the beginning of this chapter: "But such achievements [antibiotics, lights, rockets] lend authority to 'science', and science . . . is linked in the public mind with atheistic materialism.")

I think this attitude largely underestimates the significance of contemporary physicalism. What is more, it doesn't really answer the question about physicalism's sudden emergence. It is not as if the prestige of physics suddenly had a big boost in the middle of the twentieth century. I would say that physics has been pretty prestigious for about 300 years, with occasional ups and downs. Yet the philosophical physicalism we are concerned with is a distinctively late twentieth-century phenomenon.

In this chapter I want to offer a different suggestion. My explanation for the rise of physicalism will be that it follows from an argument, or rather a family of arguments, the crucial premise of which was not available, at least to philosophers, until relatively recently. This is because this crucial premise is an empirical claim, and the evidence for it has only become clear-cut over the last century. Prior to that, this premise was not upheld by scientific theory, and so was unavailable as a basis for philosophical argument.

If this explanation is right, it casts a different light on physicalist views. Physicalism has been pressed on philosophers, not by fad or fashion, but by a newly available line of argument. In saying this, I do not want to suggest that the argument for physicalism is uncontroversial, or that the crucial premise I shall focus on is incontrovertible. But I do want to urge that physicalism deserves to be taken seriously, and that those who want to oppose it have an obligation to show where the argument in its favor goes wrong.

Of course, there are those, such as Stephen Clark, who think that "no one could seriously, rationally suppose" that empirical considerations could possibly yield a disproof of mind-body dualism. I shall not explicitly engage with this attitude in what follows, but shall merely invite those who find it plausible to consider the matter again at the end of this chapter. Of course, to repeat a point just made, the empirically based arguments in favor of physicalism are not incontestable. But, even so, it scarcely follows that you have to be unserious or irrational to suppose that they in fact succeed in establishing physicalism. Indeed it is my contention in this chapter that a number of the most influential of late twentieth-century analytic philosophers have supposed just that.

Phenomenalism and Physicalism

Before I give my own explanation for the rise of physicalism, in terms of the new availability of an empirical argument, let me quickly consider an alternative possible explanation, namely, that the rise of physicalism is simply the other side of the demise of phenomenalism.

No doubt there is something to this thought. Phenomenalism was the dominant metaphysical view among logical positivists and other scientifically minded analytic philosophers in the first half of this century. And there certainly isn't much room within phenomenalism to be a physicalist. If you think that everything, including physical stuff, is logically constituted out of mental items such as sense data, then you would seem already to have ruled out the thought that mental items are in turn constituted by physical items.

Even so, I don't think this is a sufficient explanation for the rise of physicalism. For one thing, the rejection of phenomenalism doesn't yet explain the acceptance of physicalism. After all, you can deny phenomenalism without embracing physicalism. Indeed a significant number of contemporary philosophers do exactly that. These philosophers reject phenomenalism, but see no reason to privilege the physical among the different categories of things that exist, and so do not agree that everything is physically constituted.

Apart from this, there is the question of why phenomenalism died in the first place. This is, of course, a big subject, and any full answer would have to mention Wittgenstein's private language argument and Sellars's attack on givens. But I suspect that just as influential as these was the empirical argument for physicalism I am about to discuss. It is a simple argument, from uncomplicated empirical premises, and phenomenalists would have been as well placed to appreciate its force as anybody else. If there is anything to this suggestion, then it wasn't so much that physicalism happened to fill the space created when phenomenalism left the stage. Rather the argument for physicalism was itself partially responsible for the overthrow of phenomenalism.

It is high time I described this empirically based argument for physicalism. It is simple enough in outline. The crucial empirical premise is *the completeness of physics*, by which I mean that all physical effects are due to physical causes. And the argument is then simply that, if all physical effects are due to physical causes, then anything that has a physical effect must itself be physical.

The important point, for our purposes, is that the premise here, the completeness of physics, is a doctrine with a history. It was not always widely accepted. In particular, it was only after some decades of the present century that it became part of scientifically educated common sense. This in turn was because evidence favoring this thesis did not start to emerge until the mid-nineteenth century and did not become generally persuasive until much later. Once the thesis was widely accepted, however, its implications were obvious, and nearly all philosophers with some acquaintance with modern physical science became physicalists.

In the rest of this chapter I shall proceed as follows. First, in the next two sections I shall get a bit clearer about what the completeness of physics says, and how different philosophers have used it to argue for physicalism. In the following sections I then shall examine the history of this thesis, and in particular the reasons why it has come to be widely accepted nowadays, even though it wasn't always.

The Completeness of Physics and the Argument for Physicalism

Let me start by formulating a more precise version of the thesis of the completeness of physics:

> All physical effects are fully determined[2] by law by prior physical occurrences.

Note first that this thesis does not yet assert physicalism. Physicalism is the doctrine that everything, including prima facie, nonphysical stuff, is physical. But the completeness of physics doesn't itself say anything about nonphysical things. It is purely a doctrine about the structure of the physical realm. It says that, if you start with some physical effect, then you will never have to leave the realm of the physical to find a fully sufficient cause for that effect.[3]

If we want to get from the completeness of physics itself to the imperialist physicalist conclusion that *everything* is physical, we need an argument. However, the general shape of such an argument is not hard to find. As I put it in the last section, if the completeness of physics is right, and all physical effects are due to physical causes, then anything *that has a physical effect* must itself be physical. Or, to put it the other way around, if the completeness of physics is right, then there is no room left for anything nonphysical to make a difference to physical effects, so anything that does make such a difference must itself be physical.

Some version of this line of thought underlies the writings of all the philosophers who started arguing for physicalism in the 1950s and 1960s. Thus, for example, consider Smart's thought that we should identify mental states with brain states, for otherwise those mental states would be 'nomological danglers' that play no role in the explanation of behavior. Similarly, reflect on Armstrong's and Lewis's argument that, because mental states are picked out by their causal roles, including their roles as causes of behavior, and because we know that physical states play these roles, mental states must be identical with those physical states. Or, again, consider Davidson's argument that, because the only laws governing behavior are those connecting behavior with physical antecedents, mental events can only be causes of behavior if they are identical with those physical antecedents.

There is much to say about these arguments, and I shall say some of it in the following text. But the point I want to make here is that none of these arguments would seem even slightly plausible without the assumption of the completeness of physics. To see this, imagine that the completeness of physics were not true, and that some physical effects (the movements of arms, perhaps, or the firings of the motor neurons that instigate those movements) were not determined by law by prior physical causes at all, but by *sui generis* nonphysical mental causes, such as decisions, say, or exercises of will, or perhaps just pains. Then (1) contra Smart, mental states wouldn't be "nomological danglers," but directly efficacious in the production of behavior; (2) contra Armstrong and Lewis, it wouldn't necessarily be physical states that played the causal roles by which we pick out mental states, but quite possibly the *sui generis* mental states themselves; and (3) contra Davidson, it wouldn't be true that the only laws governing behavior are those connecting behavior with physical antecedents, because there would also be laws connecting behavior with mental antecedents.[4]

Comments on the Causal Argument for Physicalism

The interesting historical question, to which I shall turn shortly, is why these completeness-of-physics-based arguments started appearing when they did. But first it will be useful to clear away a bit of philosophical undergrowth. Those readers who are more interested in history than philosophical niceties may wish to skip ahead to the next section.

There are significant differences between the completeness-based arguments put forward by Smart, Armstrong, Lewis, Davidson, and other physicalist writers. However, rather than getting entangled in detailed comparisons, let us focus on one canonical form of this argument, which I shall call the 'causal argument' (Crane [1995]; Sturgeon [1998]). This will enable me to make some general structural points.

Premise 1 (the completeness of physics):

All physical effects are fully determined by law by prior physical occurrences.

Premise 2 (causal influence):

All mental occurrences have physical effects.[5]

Premise 3 (no universal overdetermination):

The physical effects of mental causes are not all overdetermined.

Conclusion:

Mental occurrences must be identical with physical occurrences.

Some comments:

(1) *The Ontology of Causes*. The force of this causal argument is extremely sensitive to how you think about causation. If, just as Donald Davidson (1980), you think of the relata of causation as events and think of events in turn as basic particulars, then the argument concludes only that mental and physical descriptions pick out the same *events*, not that there is any constitutive relationship between mental and physical *properties*. On the other hand, if you think of the relata of causation as

instantiations of properties, or more generally as facts (Mellor [1995]), then the argument promises to establish the stronger conclusion that mental properties are identical with physical ones. Because the stronger version is the more interesting, and because facts in any case seem to me the better candidates for the relata of causation, I shall read the argument in this way henceforth.

(2) *Accepting Overdetermination*. The causal argument seems pretty clearly valid.[6] So those who reject the conclusion must reject one of the three premises. All three moves are found in the literature. The status of premise 1, the completeness of physics, will occupy most of what follows. This leaves premises 2 and 3. Let us first consider rejecting premise 3, the premise of no universal overdetermination.

To reject this premise is to accept that the physical effects of mental causes are always overdetermined. This 'belt and braces' view is defended by Gabriel Segal and Elliott Sober (1991) and D.H. Mellor (1995, pp. 103–5). In response to the worry that this view seems to imply that your arm would still have moved even if you hadn't felt a pain (because your C-fibers would still have fired, say), these philosophers argue that the distinct mental and physical causes may themselves be strongly counterfactually dependent. Still, this then raises the question of *why* such causes should always be so counterfactually dependent, if they are genuinely distinct. Possible causal mechanisms underpinning this dependence can be imagined, but there seems to me no good reason to believe in them.

(3) *Epiphenomenalism and Preestablished Harmony*. What about premise 2? The possibility of denying this premise is familiar enough, under the guise of 'epiphenomenalism' or 'preestablished harmony.' If you are prepared to accept that mental states do not have physical effects, and are indeed 'nomological danglers' with respect to the causation of behavior, then the previously mentioned argument for physicalism will not move you, for you will not embrace its second premise. I leave it to readers to decide whether this denial of the efficacy of the mental is a price worth paying to avoid physicalism.[7]

While we are on this point, it is worth noting that one of the most popular versions of physicalism, namely, functionalism, is arguably a closet version of epiphenomenalism. By functionalism I mean the view that identifies a mental state with a 'second-order state,' that is, the state-of-having-some-state-that-plays-a-certain-role, rather than with the first-order physical state that actually plays that role. Because the second-order mental state cannot be *identified* with the first-order physical state (rather, it is 'realized' by it), it is not clear that it can be deemed to cause what that first-order state causes, such as items of behavior. So functionalism threatens the epiphenomenalist denial of premise 2, the claim that mental states have physical effects.

The recognition of this difficulty has put functionalism under some pressure recently. One option is to turn away from functionalism and insist that mental states are first-order states after all, and so strictly identical with physical states (Lewis [1980]). This option in effect upholds a strong version of premise 2, and allows it to argue for the full identity of mental with physical properties. Another option is to read 'causation' generously in premise 2, so as to allow that second-order states 'cause' what their realizers cause. Taken this way, the causal argument then yields the weaker conclusion that mental states must be physically realized second-order states (for, if they weren't at least this, the completeness of physics wouldn't allow them to 'cause' behavior even in the weaker sense). (For more on this issue, see Papineau [1998].)

(4) *Noncausal Realms*. This discussion of epiphenomenalism shows that the causal argument for physicalism only applies to nonphysical occurrences that *do* have physical effects. Without premise 2, there is no argument, because it is only on the assumption that the nonphysical occurrences in question are *not* 'causal danglers' that we need to identify them with something physical.

This shows that there are limits to this form of argument for physicalism. At the beginning of this chapter I characterized physicalism as the doctrine that 'everything is physically constituted.' However, this ambitious claim outstrips anything that can be delivered by the causal argument. For the causal argument has no grip on putative realms of reality that are outside the causal realm altogether, and so a fortiori don't have physical effects. I particularly have in mind here the realms of mathematics, and of moral and other values. While some philosophers have supposed that mathematical or moral facts do have physical effects, this is not the normal way to think about them. And, if we do deny that moral or mathematical facts have physical effects, then our causal argument will provide no basis for identifying them with physical facts.[8]

I myself think that this limitation to the causal argument constitutes a genuine boundary to the proper ambitions of physicalism. I think that physicalism is best formulated, not as the claim that everything is physical, but as the significantly weaker claim that everything that interacts causally with the physical world is physical. This leaves it open that there may be noncausal realms of reality that are not physically constituted, such as the realm of moral worth, or of beauty, or of mathematical objects.

Of course, there may be other problems with such nonphysical realms. For example, it is not clear how we may come by knowledge of such realms, if they can have no physical effects on our sense organs. But these further arguments are by no means clear-cut, and there is no obvious reason why they should be accepted by everybody who accepts the causal argument. Because of this, I shall use 'physicalism' in the rest of this chapter specifically for the doctrine that everything with causal powers is physical, whatever may be true of noncausal realms.

(5) *What is "Physics"?* In a moment I shall turn to the history of the completeness of physics. But first we need to address a terminological issue, one that may have been worrying readers for some time. How exactly is *physics* to be understood in this context of the causal argument? An awkward dilemma may seem to face anyone trying to defend the crucial first premise, the completeness of physics. If we take *physics* to mean the subject matter currently studied in departments of physics, discussed in physics journals, and so on, then it seems pretty obvious that physics is not complete. The track record of attempts to list *all* the fundamental forces and particles responsible for physical effects is not good, and it seems highly likely that future physics will identify new categories of physical cause. On the other hand, if we mean, by *physics*, the subject matter of such future scientific theories, then we seem to be in no position to assess its completeness, because we don't yet know what it is.

This difficulty is more apparent than real. If you want to use the causal argument, it isn't crucial that you know exactly what a complete physics would include. Much more important is to know what it won't include. Suppose, for example, that you have an initial idea of what you mean by *mental* (the sentient, say, or the intentional, or perhaps just whatever events occur specifically in the heads of intelligent beings). And suppose now that you understand *physical* as simply meaning *nonmental*. Then, provided we can be confident that the "physical" in this sense is complete, that is, that every nonmental effect is fully determined by *nonmental* antecedents (in the sense of antecedents that *can* be identified without using mental categories), then we can conclude that all mental states must be identical with something nonmental (otherwise mental states couldn't have nonmental effects). This understanding of *physical* as *nonmental* might seem a lot weaker than most pretheoretical understandings, but note that it is just what we need for philosophical purposes, because it still generates the worthwhile conclusion that the mental must be identical with the nonmental – given, that is, that we are entitled to assume that the nonmental is complete.

The same point applies if we want to apply the causal argument to chemical, biological, or economic states. As long as we can be confident that all nonchemical effects are fully caused by nonchemical (nonbiological, noneconomic . . .) states, then we can conclude that all chemical (biological, economic . . .) states must be identical with something nonchemical (nonbiological, noneconomic . . .).

We might not know enough about physics to know exactly what physics does include. But as long as we are confident that it excludes such-and-such special categories, then we can use the causal argument to conclude that these special categories are in fact identical with other kinds. I shall suppose this indirect understanding of *physics* in what follows: it should simply be understood as that set of properties that can be specified without appeal to whichever special vocabularies (mental, biological, . . .) we are interested in. Correspondingly, the completeness of physics will be the doctrine that such nonspecial effects are always fully accounted for by nonspecial causes (cf. Papineau and Spurrett [1998]).

Descartes and Leibniz

Let us now concentrate on the history of the completeness of physics. The important question, as we have just seen, is whether any nonspecial effects are produced by *sui generis* special causes. True, the exact content of this question will be relative to which special categories we are interested in, for the reasons just explained. Still, we can take it for the moment that we are interested in a relatively strong version of the completeness of physics, and in particular one that would rule out *sui generis* mental and biological causes, let alone economic, social, or other even more special causes.

When I first became interested in the causal argument a few years ago, I recognized that there were many points where it could be queried. However, I assumed that the completeness premise was quite uncontentious. Surely, I thought, everybody agrees that the movements of matter, such as the movements of molecules in your arm, can in principle always be fully accounted for in terms of prior physical causes, such as physical activity in your nerves, which in turn is due to physical activity in your brain, . . . and so on.

To my surprise, I discovered that some people didn't agree. They didn't see why some physical occurrences, in our brains perhaps, shouldn't have irreducibly mental causes. My first response, when presented with this thought, was to attribute it to an insufficient education in the physical sciences. Sometimes I went so far as to communicate this diagnosis to those who disagreed with me. However, when they then asked me, not unreasonably, to show them where the completeness of physics is written down in the physics textbooks, I found myself somewhat embarrassed. Once I was forced to defend it, I realized that the completeness of physics is by no means self-evident. Indeed further reading has led me to realize, far from being self-evident, it is an issue on which the post-Galilean scientific tradition has changed its mind several times.

My original thought was that the completeness of physics would follow from the fact that physics can be formulated in terms of conservation laws. If the laws of mechanics tell us that important physical quantities are conserved regardless of what happens, then doesn't it follow that the later states of physical systems are always fully determined by their earlier physical states?

Not necessarily. It depends on what conservation laws you are committed to. Consider Descartes's mechanics. This incorporated the conservation of what Descartes called 'quantity of motion,' by which he meant mass times speed. That is, Descartes held that the total mass times speed of any collection of bodies is guaranteed to remain constant, whatever happens to them. However, this alone does not guarantee that physics is complete. In particular, it does not rule out the possibility of physical effects that are due to irreducibly mental causes.

This is because Descartes's *quantity of motion* is a nondirectional (scalar) quantity, defined in terms of speed, as opposed to the directional (vectorial) Newtonian notion of linear *momentum*, defined in terms of velocity. Because of this, the *direction* of a body's motion can be altered without altering its quantity of motion. As Roger Woolhouse explains the point, in an excellent discussion of the relevance of seventeenth-century mechanics to the mind-brain issue (1985), a car rounding a corner at constant speed conserves its 'quantity of motion,' but not its momentum.

This creates room for *sui generis* mental causes to alter the *direction* of a body's motion without violating Descartes's conservation principle. Descartes's conservation principle does mean that, if one physical body starts going faster, this must be due to another physical body going slower. But his principle doesn't require that, if a physical body changes direction, this need result from any other physical body changing direction. Even if the change of direction results from an irreducibly mental cause, the quantity of motion of the moving body remains constant.

According to Leibniz, Descartes exploited this loophole to explain how the mind could affect the brain. As Leibniz tells the story, Descartes believed that the mind nudges moving particles of matter in the pineal gland, causing them to swerve without losing speed, like the car going round the corner. This then explained how the mind could affect the brain without violating the conservation of 'quantity of motion' (Leibniz [1898] [1696], p. 327).

Now, there is little evidence that Descartes actually saw things this way, nor indeed that he was particularly worried about how the laws of physics can be squared with mind-brain interaction. Still, whatever the truth of Leibniz's account of Cartesian theory, his next point deserves our attention. For Leibniz proceeds from his analysis of Descartes to the first-order assertion that the *correct* conservation laws, unlike Descartes's conservation of quantity of motion, *cannot* in fact be squared with mind-body interaction.

Leibniz's conservation laws were in fact a great improvement on Descartes's. In place of Descartes's conservation of 'quantity of motion,' Leibniz upheld both the conservation of linear *momentum* and the conservation of kinetic *energy*. These two laws led him to the correct analysis of impacts between moving bodies, a topic on which Descartes had gone badly astray. And, in connection with our present topic, they persuaded him that there is no room whatsoever for mental activity to influence motion of matter.

In effect, the conservation of linear momentum and of kinetic energy together squeeze the mind out of the class of events that cause changes in motion. Leibniz's two conservation laws, plus the standard seventeenth-century assumption of no physical action at a distance, are themselves sufficient to fix the evolution of all physical processes. The conservation of momentum requires the preservation of the same total amount of quantity of motion in *any given direction*, thus precluding any possibility of mental nudges altering the direction of moving physical particles. Moreover, the conservation of energy, when added to the conservation of momentum, fully fixes the speed and direction of impacting physical particles after the collide.[9] So there is no room for anything else, and in particular for anything mental, to make any difference to the motions of physical particles, if Leibniz's two conservation laws are to be respected.

We can simplify the essential point at issue here by noting that Leibniz's conservation laws, unlike Descartes's, ensure physical determinism. They imply that the physical states of any system of bodies at one time fix their state at any later time. Physical determinism in this sense is certainly sufficient for the completeness of physics, even if the possibility of quantum-mechanical indeterminism means it is not necessary (cf. Note 2). So Leibniz's dynamics, unlike Descartes's, make it impossible for anything except the physical to make a difference to anything physical.

Leibniz was fully aware of the implications of his dynamical theories for mind-body interaction (cf. Woolhouse, op. cit.). However he did not infer mind-brain identity from his commitment to the completeness of physics. Instead he adopted the doctrine of preestablished harmony, according to which the mental and physical realms are each causally closed, but prearranged by the divine will to march in step in such a way as to display the standard mind-brain correlations. In terms of the canonical causal argument laid out in the section on the causal argument for physicalism, Leibniz is denying premise 2 here, about the causal influence of mind on matter. He avoids identifying mental causes with physical causes, in the face of the completeness of physics, by denying that mental causes ever have physical effects.

Newtonian Physics

Some readers might wonder why this isn't the end of the issue. Given that Leibniz established, against Descartes, that both momentum and energy are conserved in systems of moving particles, then why wasn't the history of the mind-brain argument already over? Of course, nowadays we might not want to follow Leibniz in opting for preestablished harmony, as opposed to simply embracing mind-brain identity. But this is simply because we favor a different response to the causal argument laid out in the section on the causal argument for physicalism, not because we have any substantial premises Leibniz lacked. In particular, the crucial first premise of the causal argument, the completeness of physics, would seem already to have been available to Leibniz. So doesn't this mean that everything needed to appreciate the causal argument was already on hand in the second half of the seventeenth century, long before the rise of twentieth-century physicalism?

Well, it was – but only on the assumption Leibniz gives us the correct dynamics. However, Leibniz's physical theories were quickly eclipsed by those of Newton, and this then reopened the whole issue of the completeness of physics.

The central point here is that Newton allowed forces other than impact. Leibniz, along with Descartes and all other pre-Newtonian proponents of the 'mechanical philosophy,' took it as given that all physical action is by contact. They assumed that the only possible cause of a change in a physical body's motion is the impact of another physical body. (Or more precisely, as we are telling the story, Descartes supposed that the only possible *nonmental* cause of physical change is impact, and

Leibniz then argued that *mental* causes other than impact are not possible either, if the conservation of momentum and energy are to be respected.)

Newtonian mechanics changed the whole picture. This is because Newton did not take impact as his basic model of dynamic action. Rather his basic notion is that of an *impressed force*. Rather than thinking of 'force' as something inside a body that might be transferred to other bodies in impact, as did all his contemporaries (and indeed as did most of his successors for at least a century[10]), Newton thought of forces as disembodied entities, acting on the affected body from outside. An impressed force "consists in the action only, and remains no longer in the body when the action is over." Moreover, "impressed forces are of different origins, as from percussion, from pressure, from centripetal force" (Newton [1960] [1686], Definition IV). Gravity was the paradigm. True, the force of gravity always arose from the presence of massive bodies, but it pervaded space, waiting to act on anything that might be there, so to speak, with a strength as specified by the inverse square law.

Once disembodied gravity was allowed as a force distinct from the action of impact, then there was no principled barrier to other similarly disembodied special forces, such as chemical forces, or magnetic forces, or forces of cohesion (cf. Newton [1952] [1704], Queries 29–31) – or indeed vital and mental forces.

Nothing in classical Newtonian thinking rules out special mental forces. While Newton has a general law about the effects of his forces (they cause proportional changes in the velocities of the bodies they act on), there is no corresponding general principle about the causes of such forces. True, gravity in particular is governed by the inverse square law, which fixes gravitational forces as a function of the location of bodies with mass. But there is no overarching principle dictating how forces in general arise. This opens up the possibility that there may be *sui generis* mental forces, which would mean that Newtonian physics, unlike Leibnizian physics, is not physically complete. Some physical processes could have nonphysical mental forces among their causal antecedents.[11]

The switch from a pure impact-based mechanical philosophy to the more liberal world of Newtonian forces thus undermined Leibniz's argument for the completeness of physics. Leibniz could hold that the principles governing the physical world leave no room for mental acts to make a difference because he had a simple mechanical picture of the physical world. Bodies preserve their motion in any given direction until they collide, and then they obey the laws of perfect elastic impact. The Newtonian picture is far less pristine, and gives no immediate reason to view physics as complete.

You might think that the conservation laws of Newtonian physics would themselves place constraints on the generation of forces, in such a way as to restore the completeness of physics. But this would be a somewhat anachronistic thought. Conservation laws did not play a central role in Newtonian thinking, at least not in that of Newton himself and his immediate followers. True, Newton's mechanics does imply the conservation of *momentum*. This falls straight out of his Third Law, which requires that 'action and reaction' are always equal. But it is a striking feature of Newtonian dynamics that there is no corresponding law for energy.[12]

Of course, as we shall see in the next section, the principle of the conservation of kinetic *and* potential energy in all physical processes did *eventually* become part of the Newtonian tradition, and this does impose a general restriction on possible forces, a restriction expressed by the requirement that all forces should be 'conservative.' But this came much later, in the middle of the nineteenth century, and so had no influence on the range of possible forces admitted by seventeenth- or eighteenth-century Newtonians. (Moreover, it is a nice question, to which we shall return at length in the following text, how far the principle of the conservation of kinetic plus potential energy, with its attendant requirement that all forces be conservative, does indeed constitute evidence against *sui generis* mental forces.)

In any case, whatever the significance of later Newtonian derivations of the conservation of energy, early Newtonians certainly saw no barrier to the postulation of *sui generis* mental forces. It will be helpful to distinguish in the abstract two ways in which such a Newtonian violation of the completeness of physics could occur.

First, and most obviously, it could follow from the postulation of *indeterministic* mental forces. If the determinations of the self (or of the 'soul,' as they would have said in the seventeenth and eighteenth centuries) could influence the movements of matter in spontaneous ways, then the world of physical causes and effects would obviously not be causally closed, because these spontaneous mental causes would make a difference to the unfolding of certain physical processes.

But, second, it is not even necessary for the violation of completeness that such *sui generis* special forces operate indeterministically. Suppose that the operation of mental forces were governed by fully *deterministic* force laws (suppose, for example, that mental forces obeyed some inverse square law involving the presence of certain particles in the brain). Then mental forces would be part of Newtonian dynamics in just the same sense as gravitational or electrical forces: we could imagine a system of particles evolving deterministically under the influence of all these forces, including mental forces, with the forces exerted at any place and time being deterministically fixed by the relevant force laws. Even so, this deterministic model would still constitute a violation of the completeness of physics, for the physical positions of the particles would depend inter alia on prior mental causes, and not exclusively on prior physical causes.

Did I not say at the end of the last section that determinism is sufficient for the completeness of physics (even if not necessary, because of quantum mechanics)? No. What I said was that *physical* determinism (the doctrine that prior *physical* conditions alone are enough to determine later physical conditions) is sufficient for the completeness of physics. However, we can accept determinism as such without accepting physical determinism, and so without accepting the completeness of physics. In particular, we can have a deterministic model in which mental forces play an essential role, and in which the physical subpart is therefore not causally closed.

You might feel (indeed might have been feeling for some time) that a realm of deterministic mental forces would scarcely be worth distinguishing from the general run of physical forces, given that they would lack the spontaneity and creativity that is normally held to distinguish the mental from the physical. And you might think that it is therefore somewhat odd to view them as violating the completeness of physics. I happily concede that there is something to this thought. But I would still like to stick to my terminology, as stipulated at the end of the section on the causal argument for physicalism, which assumed an initial sense for *mental* (as sentient, intentional, or intelligent), and then defined the *physical* as whatever can be identified without alluding to such mental properties – which then makes even deterministically governed *sui generis* mental forces come out 'nonphysical,' because they can't be so nonmentally identified. This is the terminology that best fits with our original interest in the causal argument for physicalism. We don't want deterministic mental forces to be counted as consistent with the completeness of 'physics,' precisely because we wouldn't be able to use *this* kind of completeness of 'physics' to infer that these mental forces are always identical with some *other* (nonmental) causes of their effects.

So far I have merely presented the possibility of special Newtonian forces as an abstract possibility. However, the postulation of such forces was commonplace among eighteenth-century thinkers, particularly among those working in anatomy and physiology. Many of the theoretical debates in these areas were concerned with the existence of vital and mental forces, and with the relation between them. Among those who debated these issues, we can find both the indeterministic and deterministic models of mental forces.[13]

Thus consider the debate among eighteenth-century physiologists about the relative roles of the forces of *sensibility* and *irritability*. This terminology was introduced by the leading German physiologist Albrecht von Haller, Professor of Anatomy at Göttingen from 1736. Haller thought of 'sensibility' as a distinctively mental force. 'Irritability' was a nonmental but still peculiarly biological power. ("What should hinder us from granting irritability to be a property of the animal *gluten*, the same as we acknowledge gravity and attraction to be properties of matter in general . . . ," Haller [1936] [1751].) Haller took the force of sensibility to be under the control of the soul and to operate solely through the nerves. Irritability, by contrast, he took to be located solely in the muscle fibers.

In distinguishing the mentally directed force of sensibility from the more automatic force of irritability, Haller can here be seen as conforming to my model of *indeterministic* mental forces. Where the force of irritability is determined by prior stimuli and is independent of mental agency, the force of sensibility responds to the spontaneous commands of the soul.

Haller's model was opposed by Robert Whytt (1714–66) in Edinburgh. In effect Whytt can be seen as merging Haller's distinct mental and vital forces, irritability and sensibility. On the one hand, Whytt gave greater power to the soul: he took it that a soul, or 'sentient principle,' is distributed throughout the body, not just in the nerves, and is responsible for all bodily activities, from the flow of blood and motion of muscles, to imagination and reasoning in the brain. But at the same time as giving greater power to this sentient principle, he also rendered its operations *deterministic*. He explicitly likened the sentient principle to the Newtonian force of gravity, and viewed it as a necessary principle that acts according to strict laws. Whytt can thus be seen as exemplifying my model of deterministic mental forces: the sentient principle is simply another deterministic Newtonian force, just like gravity and the others, in that its operations are fixed by a definite force law (Whytt [1755]).

The Conservation of Energy

In this section I want to consider how the principle of the conservation of energy eventually emerged within the tradition of Newtonian mechanics, and how this bears on the completeness of physics. It will be useful to separate some different aspects of this emergence.

Rational Mechanics

Through the eighteenth and early nineteenth centuries a number of mathematician-physicists, among the most important of whom were Jean d'Alembert (1717–83), Joseph Louis Lagrange (1736–1813), the Marquis de Laplace (1749–1827), and William Hamilton (1805–65), developed a series of mathematical frameworks designed to simplify the analysis of the motion of interacting particles. These frameworks allowed physicists to abstract away from detailed forces of constraint, such as the forces holding rigid bodies together, or the forces constraining particles to move on surfaces, and concentrate on the effects produced by other forces. (See Elkana [1974], ch. II, for the history, and Goldstein [1964], for the mathematics.)

These mathematical developments also implied that, under certain conditions, the sum of kinetic energy and potential energy remains constant. Roughly, when all forces involved are independent of the velocities of the interacting particles and of the time (let us call forces of these kinds *conservative*), then the sum of actual kinetic energy (measured by $\frac{1}{2}mv^2$) plus the potential to generate more such energy (often called the 'tensions' of the system) is conserved: when the particles slow down, this builds up 'tensions,' and, if those 'tensions' are expended, the particles will speed up again.

We now think of this as the most basic of all natural laws. But this attitude was not part of the original tradition in rational mechanics. There were two reasons for this. First, the Newtonian scientists in this tradition were not looking for conserved quantities anyway. As I explained earlier, conservation principles played little role in classical Newtonian thinking. True, Leibniz had urged the conservation of kinetic energy (under the guise of 'vis viva'), but by the eighteenth century Leibniz's influence had been largely eclipsed by Newton's. Second, the conservation of potential and kinetic energy in any case only holds under the assumption that all forces are conservative. Nowadays we take this requirement to be satisfied for all fundamental forces. But this again was no part of eighteenth-century thinking. Some familiar forces happen to be conservative, but plenty of other forces are not. Gravitation, say, is conservative, because it depends only on the positions of the particles, and not on their velocities, nor on the elapsed time. But, by contrast, frictional forces are not conservative, because they depend on the velocity of the decelerated body relative to the medium. And correspondingly frictional forces do not in any sense seem to conserve energy: when they decelerate a body, no 'tension' is apparently built up waiting to accelerate the body again.

For both these reasons, the tradition in rational mechanics did not initially view the conservation of kinetic and potential energy in certain systems as of any great significance. On the contrary, it was simply a handy mathematical consequence that falls out of the equations when the operative forces all happen to fall within a subset of possible forces (cf. Elkana [1974], ch. 2).

Equivalence of Heat and Mechanical Energy

In the first half of the nineteenth century a number of scientists, most prominently James Joule (1819–89), established the equivalence of heat and mechanical energy, in the sense of showing that a specific amount of heat will always be produced by the expenditure of a given amount of mechanical energy (as when a gas is compressed), and vice versa (as when a hot gas drives a piston).

These experiments suggested directly that some single quantity is preserved through a number of different natural interactions. They also had a less direct bearing on the eventual formulation of the conservation of energy. They indicated that apparently nonconservative forces such as friction and other dissipative forces need not be nonconservative after all, because the kinetic energy apparently lost when they acted could in fact be preserved by the heat energy gained by the resisting medium.[14]

The stage was now set for the formulation of a universal principle of the conservation of energy. We can distinguish three elements that together contributed to the formulation of this principle. First, the tradition of rational mechanics provided the mathematical scaffolding. Second, the experiments of Joule and others suggested that different natural processes all involve a single underlying quantity that could manifest itself in different forms. Third, these experiments also suggested that apparently nonconservative forces such as friction were merely macroscopic manifestations of more fundamental conservative forces.

Of course, it is only with the wisdom of hindsight that we can see these different strands as waiting to be pulled together. At the time, they were hidden in abstract realms of disparate branches of science. It took the genius of the young Hermann von Helmholtz (1821–94) to see the connections. In 1847, at the age of twenty six, he published his monograph *Uber die Erhaltung der Kraft* (*On the Conservation of Force*). The first three sections of this treatise are devoted to the tradition of rational mechanics, and in particular to explaining how the total mechanical energy (kinetic plus potential energy) in a system of interacting particles is constant in those cases where all forces are familiar 'central forces' independent of time and velocity. The fourth section describes the equivalence between mechanical 'force' and heat, referring to Joule's results, while the last two sections extend the discussion to electric and magnetic 'forces,' showing again that there are fixed equivalences between these 'forces,' heat, and mechanical energy.[15]

Physiology

At the end of his treatise Helmholtz touches on the conservation of energy in living systems. Helmholtz was in fact a medical doctor by training, and had been a student in the Berlin physiological laboratory of Johannes Müller in the early 1840s, along with Emil Du Bois-Reymond (1818–96) and Ernst Brücke (1819–92). Together these students were committed to a reductionist program in physiology, aiming to show that phenomena such as respiration, animal heat, and locomotion could all be understood to be governed by the same laws as operate in the inorganic realm.

This physiological context undoubtedly played a fundamental role in Helmholtz's articulation of a universal principle of the conservation of energy. Because of his physiological interests, Helmholtz was interested in a principle that would cover *all* natural phenomena, including those in living systems, and not just such manifestly physical phenomena as mechanical motion, heat, and electromagnetism. Thus he took the crucial step of asserting that *all* forces conserve the sum of kinetic and potential energy. Superficially nonconservative forces such as friction are simply macroscopic manifestations of more fundamental forces that preserve energy at the microlevel. This then enabled Helmholtz to view the equivalences established by experimentalists such as Joule, not just as striking local regularities, but as necessary consequences of a fundamental principle of mechanics. All natural processes must respect the conservation of energy, including processes in living systems.

It is noteworthy that neither the experimentalists such as Joule, nor the mathematician-physicists in the rational mechanics tradition, made this crucial step to a universal principle. None of the scientists working experimentally on numerical equivalences between different processes, such as Joule, generalized their discoveries into the claim that there is one quantity, energy, preserved in all natural interactions whatsoever. While it is true that a number of different scientists at the time were investigating such numerical equivalences (thus the historical thesis of the 'simultaneous discovery' of the conservation of energy), there is no reason to suppose that these scientists were generally inspired by any vision of the underlying unity of different natural processes. Similarly, there was nothing to attract mathematical physicists in the tradition of rational mechanics to the conclusion that all forces are conservative, for the reasons given previously. They simply thought of such forces as the mathematically tractable special case where changes in kinetic and potential energy always happen to balance out.

Without the desire to bring living systems under a unified science, none of these scientists had any motive for synthesizing the different strands pulled together by Helmholtz. It was Helmholtz's combination of physiological interests and sophisticated physical understanding that precipitated the crucial step. He saw that, if we assume that all fundamental forces are conservative, then this guarantees that a certain quantity, the total energy, will be preserved in all natural processes whatsoever, including the organic processes that formed the focus of his interest.

Vital Forces

Helmholtz was part of a tradition in experimental physiology that set itself in opposition to the previous generation of German *Naturphilosophen*. These nature-philosophers had developed the eighteenth-century notions of 'irritability' and 'sensibility' into a philosophy of *Lebenskraft*. By the first half of the nineteenth century, this notion of vital force had broken loose from its original Newtonian moorings and had become part of a florid metaphysics imbued with romanticism and idealism. According to the *Naturphilosophen*, organic matter was imbued with a special power, the *Lebenskraft*, which organized and directed it. They viewed this power as having a quasi-mental aspect, which enabled it to mediate between the 'archetypical ideas' or 'essences' of different species and the development of individual organisms toward that ideal form (see Coleman [1971], ch. 3; Steigerwald [1998]).

The experimental tradition that included Helmholtz can be seen as a reaction to these extreme doctrines. However, it is striking that many of those associated with this tradition, though not Helmholtz himself, continued to admit the possible existence of vital forces, both before and after the emergence of the conservation of energy. This is less puzzling than it may at first seem. These physiological thinkers did not think of vital forces as the mystical intermediaries of the *Naturphilosophen*, imbued with all the powers of creative mentality. Rather these thinkers were reverting to the tradition of eighteenth-century physiology. They simply viewed vital forces as special Newtonian forces, additional to gravitational forces, chemical forces, and so on, and which happen to arise specifically in organic contexts. Justus von Leibig (1803–73), the leading physiological chemist of the time, and Müller, Helmholtz's own mentor, are clear examples of experimental physiologists who were prepared to countenance vital forces in this sense (cf. Coleman [1971], ch. VI; Elkana [1974], ch. IV).

Does the Conservation of Energy Rule out Vital (and Mental) Forces?

The interesting question, from the point of view of this chapter, is how far this continuing commitment to vital forces is consistent with the doctrine of the conservation of energy. There is certainly some tension between the two doctrines. It is noteworthy that Helmholtz, and his young colleagues from Müller's laboratory, were committed to the view that no forces operated inside living bodies that were not also found in simpler physical and chemical contexts (Coleman [1971], pp. 150–4). Even so, there is no outright inconsistency between the conservation of energy and vital forces; and many late nineteenth-century figures were quite explicit, not to say enthusiastic, about accepting both.

In order to get clearer about the room left for vital (or mental) forces by the conservation of energy, recall how earlier I distinguished two ways in which early Newtonian theory left room for such special forces to violate the completeness of physics. First, such forces might operate spontaneously and indeterministically: nothing in early Newtonian theory would seem to rule out spontaneous forces ungoverned by any deterministic force law. Second, even if the relevant forces are governed by a deterministic force law, they may still be *sui generis*, in the sense that they may be distinct from gravitational forces, chemical forces, and so on, and may arise specifically in living systems or their brains.

The conservation of energy bears differentially on these two kinds of special forces. It does seem inconsistent with the first kind of special force, a spontaneous special force. But it does not directly rule out the second, deterministic kind.

Why should the conservation of energy rule out even a spontaneous special force? (Think of a spontaneous mental force that accelerates molecules in the pineal gland.) Why shouldn't such a force simply respect the conservation of energy, by not causing accelerations that will violate it? But this doesn't really make sense. The content of the principle of the conservation of energy is that losses of kinetic energy are compensated by build-ups of potential energy, and vice versa. But we couldn't really speak of a 'build-up' or 'loss' in the potential energy associated with a force, if there were no force law governing the deployment of that force. So the very idea of potential energy commits us to a law that governs how the relevant force will cause accelerations in the future.

However, nothing in this argument rules out the possibility of vital, mental, or other special forces that *are* governed by deterministic force laws. After all, the conservation of energy in itself does not tell which basic forces operate in the physical universe. Are gravity and impact the only basic forces? What about electromagnetism? Nuclear forces? And so on. Clearly the conservation of energy as such leaves it open as to exactly which basic forces exist. It only requires that, whatever they are, they operate conservatively.

The Death of Emergentism

So a commitment to the conservation of energy by no means settled the question of whether *sui generis* mental or vital forces should be rejected and physics declared complete. True, some few thinkers, such as Helmholtz himself, conjoined the conservation of energy with a denial of such special forces. But this was by no means mandated by the conservation of energy itself, for the reasons I have just explained. Accordingly, many other thinkers in the late nineteenth and early twentieth centuries took the opportunity to posit special forces of the kind allowed by the conservation of energy. So I still owe an explanation of what finally created a scientifically informed consensus against such special forces.

The issue is not straightforward, and there is no question of dealing with it fully here. But in this final section I would like to offer some outline conjectures. I shall proceed as follows. First, I shall take it as given that the conservation of energy at least was a settled doctrine. Of course there is a story to be told about this as well. But, for whatever reasons, the doctrine of the conservation of energy did win widespread acceptance within a decade or two of its initial formulation, and certainly none of the developments I am about to consider questioned its validity. Second, I shall lump mental and vital forces together. There are, of course, considerations that bear differentially on the existence of such forces, but I shall be proceeding at a level where these are not significant.

Two Arguments

My central suggestion will be that two rather different lines of evidence contributed to the demise of special forces. The first was an abstract argument based on theoretical physics, while the second was a more direct empirical argument based on physiological research. The abstract argument involves considerations to do with the conservation of energy, and was available from the time of Helmholtz onward (even though it was not incontrovertible, and many were not persuaded). By contrast, the direct argument has little to do with the conservation of energy, and indeed did not really gain force until the twentieth century.

At the end I shall argue that both arguments can be seen as contributing to the general modern acceptance of the completeness of physics. But the precise timing of this acceptance, and in particular the arrival of a general consensus in the second half of the twentieth century, seems to call for explanation in terms of the build-up of direct evidence for the second argument, rather than in terms of the more abstract argument that had been available since the middle of the nineteenth century.

Let me begin by presenting the two arguments in outline.

(1) *The Argument from Fundamental Forces*. The first argument is that all apparently special forces characteristically *reduce* to a small stock of basic physical forces that conserve energy. Causes of macroscopic accelerations standardly turn out to be composed out of a few fundamental physical forces that operate throughout nature. So, while we ordinarily attribute certain physical effects to 'muscular forces,' say, or indeed to 'mental causes,' we should recognize that these causes, just as all causes of physical effects, are ultimately composed of the few basic physical forces.

(2) *The Argument from Physiology*. The second argument is simply that there is no direct evidence for vital or mental forces. Physiological research reveals no phenomena in living bodies that manifest such forces. All organic processes in living bodies seem to be fully accounted for by normal physical forces.

I take both of these to be empirically based arguments, and both to have the same conclusion: namely, that there are no special mental or vital forces. But note that the evidential basis for the two arguments is quite different. The second argument appeals directly to the evidence uncovered by physiological research. It notes that observations made inside living bodies never reveal any accelerations that cannot be attributed to normal physical forces. The first argument, by contrast, appeals to the investigation of forces in general. It rests on evidence that many apparently different kinds of forces turn out to be composed of a few fundamental forces, and then applies this lesson to vital and mental forces in particular. So it need not appeal directly to any evidence about what goes on in living bodies. Instead it can infer the general conclusion inductively from the study of other forces, and then project it to the special case of mental and vital forces.

The Argument from Fundamental Forces

Let me now explain the first argument more fully. I shall return to the second argument in the following text. I take the materials for the first argument to have been available from the middle of the nineteenth century, and to relate to the reasoning that led up to the acceptance of the conservation of energy. It is true, as I have stressed, that the doctrine of the conservation of energy is itself consistent with the existence of special forces, as long as those forces are conservative. At the same time, it seems to me that the thinking that supported the conservation of energy also weighed against special mental or vital forces.

At its simplest, my thought here is that the arguments behind the conservation of energy give inductive reason to suppose that all forces reduce to a small number of fundamental forces. We have already seen how Helmholtz's formulation of the conservation of energy hinged on the assumption that friction and other dissipative forces are nonfundamental forces, macroscopic manifestations of processes involving more fundamental conservative forces. For it is only if we see macroscopic forces such as friction as reducing to fundamental conservative forces that we can uphold the universal conservation of energy. Now, this point can be viewed as providing inductive support for the general thesis that *all* apparently special forces reduce to a small stock of fundamental forces. The special forces that have been quantitatively analyzed, such as friction, turn out to reduce to more fundamental conservative forces. So this provides inductive reason to conclude that any other apparently special forces, such as muscular forces, vital forces, or mental forces, will similarly reduce.

This is of course not a knock-down argument. Vital or mental forces could figure among the fundamental forces of nature, even if they are only generated in the special circumstances associated with life or sentience. But this position does not sit happily with a continued commitment to the universal conservation of energy. An insistence on the independent existence of *sui generis* special forces inside bodies threatens to remove the reasons for believing in the conservation of energy in the first place. For there are no obvious grounds for expecting such *sui generis* special forces to be conservative.

After all, what argument was there, in 1850, say, for believing that forces operating inside bodies do not violate the conservation of energy? I am suggesting that the most persuasive argument hinged on the assumption that all forces operating in special circumstances reduce to a small stock of fundamental conservative forces. However, suppose now that it is explicitly specified that vital and mental forces do *not* reduce to other forces. Now we need independent evidence for supposing they are conservative, and it is not clear where it is to come from. In effect, then, positing *sui generis* vital or mental forces threatens to undermine the inductive grounds for upholding the conservation of energy in the first place. For it makes the assumption of their conservativeness an independent assumption, an assumption for which we lack any independent evidence.

I suspect that something like this line of thought lay behind Helmholtz's and his younger contemporaries' conviction that there were no special vital forces. Consider how Helmholtz argues in *Uber die Erhaltung der Kraft*. He takes pains to stress it is specifically *central* forces independent of time and velocity that ensure the conservation of energy. This emphasis on central forces (by which Helmholtz meant forces that act along the line between the interacting particles) now seems dated. Nowadays, conservativeness is normally defined circularly, as a property of those forces that do no work round a closed orbit, and that are therefore the gradient of a scalar that depends only on position. This definition does not require a restriction to central forces. However, Helmholtz was in no position to adopt our circular definition of conservativeness. He was aiming to *persuade* his readers of the general conservation of energy, and so needed an argument. It wouldn't have served simply to observe that energy is conserved by those forces that conserve energy. Helmholtz's actual claim was that energy is conserved by a wide range of known forces, namely, central forces. Still, this by itself doesn't show energy is conserved by all forces, *unless* all forces are central. Why should this be? Well, as previously mentioned, the most plausible thought is surely that there is a small stock of basic central forces, and that all causes apparently peculiar to special circumstances are composed out of these.

Even this is scarcely conclusive. Those thinkers who remained convinced, for whatever reasons, that there must be irreducible special forces inside living bodies, could still respect the universal conservation of energy, by maintaining that these extra forces must themselves operate conservatively. In support of this they could have offered the alternative inductive argument that, because all the *other* fundamental forces examined so far have turned out to be conservative, we should infer that any extra vital or mental fundamental forces will be conservative too.

I am not sure to what degree these alternative lines of inductive reasoning can be found explicitly laid out in the nineteenth-century debates. But they offer one possible explanation for the two different views on *sui generis* special forces that coexisted after the emergence of the conservation of energy. The thought that all apparently special forces reduce to a small stock of fundamental forces can account for the rejection of irreducible vital or mental forces by thinkers such as Helmholtz and his young colleagues. Yet there were at least as many who wanted to maintain that vital and mental forces are *sui generis*, and they had the option of arguing that, even if these forces are fundamental and irreducible, the nature of other fundamental forces provides inductive reason to suppose these *sui generis* forces will be conservative in their own right.

In connection with this latter school of thought, I have already mentioned Leibig and Müller, two eminent physiologists of the older generation, who continued to accept vital forces, even after the conservation of energy had won general acceptance. Also, Brian McLaughlin, in his excellent article on "British Emergentism" (1992), explains how the philosophers J.S. Mill and Alexander Bain went so far as to argue that the conservation of energy, and in particular the notion of potential energy, lends definite support to the possibility of nonphysical forces. (The 'British Emergentists' discussed by McLaughlin were a philosophical movement committed precisely to nonphysical causes of motion in my sense, causes that were not the vectorial "resultants" of basic physical forces such as gravity and impact, but which "emerged" when matter arranged itself in special ways. The particular idea that attracted Mill and Bain was that these "emergent forces" might be stored as unrealized potentials, ready to manifest themselves as a cause of motion only when the relevant special circumstances arose.[16,17])

The Argument from Physiology

McLaughlin explains how 'British Emergentism' continued to flourish into the twentieth century.[18] This highlights the question with which I began this chapter. Given that thinkers continued to posit special mental and vital forces until well after the Great War, why has the idea of such forces now finally fallen into general disfavor?

Here I think we need to refer to the second line of argument against such forces, the argument from direct physiological evidence. We can view this second argument as operating against the background provided by the earlier argument from fundamental forces. The earlier argument suggested that at least most natural phenomena, if not all, can be explained by a few fundamental physical forces. This focused the issue of what kind of evidence would demonstrate the existence of extra mental or vital forces. For once we know that other forces exist, then we will know which anomalous accelerations would indicate the presence of special mental or vital forces. Against this background, the argument from physiology is then simply that detailed modern research has failed to uncover any such anomalous physical processes.

The relevant research dates mostly from the twentieth century. While important physiological research was carried out in the second half of the nineteenth century (see Coleman [1971]), it did not penetrate to the level of forces operating inside bodies. At most it identified the chemical inputs and outputs to various parts of the body and showed that animals are subject to general conservation principles. (See in particular Coleman, pp. 140–3, for Max Rubner's elaborate 1889 respiration calorimeter experiments, showing that the energy emitted by a small dog exactly corresponds to that of the food it consumes.) Experiments of this kind, however, failed to provide compelling evidence against vital or mental forces. That normal chemicals are moved around, and that energy is conserved throughout, does not in the end rule out the possibility that some accelerations within bodies are due to special vital or mental forces. It may still be that such forces are activated inside cells, but operate in such a way as to "pay back" all the energy they "borrow," and vice versa.[19]

In the first half of the twentieth century the situation changed, and by the 1950s it had become difficult, even for those who were not moved by the abstract argument from general reducibility, to continue to uphold special vital or mental forces. A great deal became known about biochemical and neurophysiological processes, especially at the level of the cell, and none of it gave any evidence for the existence of special forces not found elsewhere in nature.

During the first half of the century the catalytic role and protein constitution of enzymes were recognized, basic biochemical cycles were identified, and the structure of proteins analyzed, culminating in the discovery of DNA. In the same period, neurophysiological research mapped the body's neuronal network and analyzed the electrical mechanisms responsible for neuronal activity. Together, these developments made it difficult to go on maintaining that special forces operate inside living bodies. If there were such forces, they could be expected to display some manifestation of their presence. But detailed physiological investigation failed to uncover evidence of anything except familiar physical forces.

In this way, the argument from physiology can be viewed as clinching the case for completeness of physics, against the background provided by the argument from fundamental forces. One virtue of this explanation in terms of two interrelated arguments is that it yields a natural explanation for the slow advance of the completeness of physics through the century from the 1850s to the 1950s. Suppose that we rank different thinkers through this period in terms of how much specifically physiological evidence was needed to persuade them of completeness, in addition to the abstract argument from fundamental forces. Helmholtz and his colleagues would be at one extreme, in deciding for completeness on the basis of the abstract argument alone, without any physiological evidence. In the middle would be those thinkers who waited for a while, but converted once initial physiological research in the first decades of this century gave no indication of any forces beyond fundamental forces found throughout nature. At the other end would be those who needed a great deal of negative physiological evidence before giving up on special forces. The existence of this spectrum would thus explain why there was a gradual build-up of support for the completeness of physics as the physiological evidence accumulated, culminating, I would contend, in a general scientific consensus by the 1950s.[20]

Conclusion

The problem I set myself at the beginning of this chapter was to explain the rise of physicalist doctrines in the second half of this century. My argument has been that this is due to contemporary agreement on the completeness of physics. In the main body of this chapter I have sought to show that this consensus is not just a fad, but a reflection of developments in empirical theory. Though it has not always been so, there is now good reason to believe the empirical thesis that all physical effects are due to physical causes. In particular, by the 1950s, there was enough physiological evidence to persuade even those scientists who were unmoved by the abstract argument from fundamental forces.

The rise of physicalism among philosophers can be seen as a reflection of this development within science. Without the completeness of physics, there is no compelling reason to identify the mind with the brain. But once the completeness of physics became part of established science, scientifically informed philosophers realized that this crucial premise could be slotted into the various alternative versions of the causal argument for physicalism. There seems no reason to look any further to explain the widespread philosophical acceptance of physicalism since the 1950s.

Of course, as with all empirical matters, there is nothing certain here. There is no knock-down argument for the completeness of physics. You could in principle accept the rest of modern physical theory, and yet continue to insist on special mental forces, which operate in as yet undetected ways in the interstices of intelligent brains. And indeed there do exist bitter-enders of just this kind, who continue to hold out for special mental causes, even after another half-century of ever more detailed molecular biology has been added to the inductive evidence that initially created a scientific consensus on completeness in the 1950s. Perhaps it is this possibility that Stephen Clark has in mind when he doubts whether any empirical considerations can "disprove" mind-body dualism. If so, there is no more I can do to persuade him of the completeness of physics. However, I see no virtue in philosophers refusing to accept a premise that, by any normal inductive standards, has been fully established by over a century of empirical research.[21]

NOTES

1. Although see pages 11–12 in the following text for some necessary qualifications.
2. Or, even more precisely, to accommodate quantum mechanical indeterminism: the *chances* of all physical occurrences are fully determined by prior physical occurrences. I shall ignore this qualification in nearly all that follows, because it would only complicate the issues unnecessarily.
3. Note, however, that while this is just a doctrine about physics, it does implicitly distinguish physics from other realms, because most other realms manifestly aren't complete in this sense. The mental isn't complete, for example, because there is no mental cause for the pain I feel when I sit on a drawing pin. Nor is the economic, because there is no economic cause for the economic costs occasioned by a hurricane. (This is why we don't find arguments aiming to show that everything is mental, or economic, parallel to the completeness-based argument that everything is physical.)
4. In other writings, the relevance of the completeness of physics does not need to be excavated, because it lies on the surface. Thus see Feigl (1958); Oppenheim and Putnam (1958).
5. Equally: all chemical/biological/social occurrences have physical effects. The causal argument provides a schema that delivers physicalism for other special subjects as readily as for the mental. In the historical discussion in the following text, various special categories will be at issue at different points. But it will often be expositorily convenient to let the mental stand for the other cases, especially when addressing issues of argumentative structure rather than historical substance. The context should make it clear when the category of the mental is so being used.
6. However Sturgeon (1998) argues that an equivocation between a quantum-theoretical sense of *physical* (in premise 1) and an everyday sense (in premise 2) invalidates the argument. This raises a number of interesting issues that I shall not be able to discuss here. But see Noordhof (1999) and Witmer (forthcoming).
7. Of course, many philosophers are moved to pay this price because they cannot believe that *conscious* occurrences in particular can be identical with physical occurrences. I do not think that this is a good motivation. However, I do accept that physicalists owe some explanation of why conscious occurrences *seem* so very different from physical ones, if they aren't. See Papineau (1993, ch. 4); (1998).

8. Conversely, those philosophers who do think that mathematical or moral facts have physical effects (in our brains, say) will come under pressure from the causal argument to identify them with physical facts.
9. Leibniz took it that all basic material particles are perfectly elastic, and that no kinetic energy is lost when they collide. He explained the apparent loss of kinetic energy when inelastic *macroscopic* bodies collide by positing increased motion in the microscopic parts of those bodies. (Thus he explains, in the fifth paper of the Leibniz-Clarke Correspondence, H. Alexander, ed. [1956]: "The author objects, that two soft or un-elastic bodies meeting together, lose some of their force. I answer, no. 'Tis true, their wholes lose it with respect to their total motion; but their parts receive it, being shaken (internally) by the force of their concourse.")
10. Cf. Papineau (1977).
11. Throughout the rest of this chapter I shall talk in terms of 'forces.' However, the issues will arise in just the same way if you regard forces as otiose, and instead think of the circumstances that "cause forces" as themselves directly causing the resulting accelerations. In that case, you will replace the question of whether there are 'mental forces' with the question of whether specifically mental initial conditions (conditions of sentience, intentionality, or intelligence, depending on how you wish initially to pick out the mental) make a difference to accelerations, in the sense of entering as antecedents into special laws about accelerations that do not follow from other laws about accelerations. More simply, are there special accelerations in brains that aren't predicted by other laws about accelerations? (Cf. McLaughlin [1992], pp. 64–5.)
12. One barrier to the formulation of an energy conservation principle by early Newtonians was their lack of a notion of potential energy, the energy "stored up" after a spring has been extended or compressed, or as two gravitating bodies move apart. Given this, there was no obvious sense in that they could view two gravitating bodies, for example, as conserving energy while they moved apart: after all, the sum of their kinetic energies would not be constant, but unequivocally decreasing. And even in the case of impact, where the notion of potential energy is not immediately needed, early Newtonians displayed no commitment to the conservation of (kinetic) energy. Most obviously, Newton and his followers were perfectly happy, unlike Leibniz, to allow unreduced inelastic collisions, in which both bodies lose kinetic energy without transmitting it to their internal parts. It is also worth remarking that there is nothing in Newton's Laws of Motion to rule out even 'superelastic' impacts, in which total kinetic energy increases. If two bodies with equal masses and equal but opposite speeds both rebounded after collision with double their speeds, for example, Newton's three Laws of Motion and the conservation of momentum would be respected. True, any such phenomenon would provide an obvious recipe for perpetual motion, but the point remains that Newton's Laws do not rule it out. (It is also worth noting that perpetual motion was by no means universally rejected by seventeenth- and eighteenth-century physicists. Cf. Elkana [1974], pp. 28–30.)
13. Here I am closely following Steigerwald (1998, ch. 2).
14. One model for this preservation was the kinetic theory of heat. This account took the macroscopic kinetic energy that was apparently lost to be converted into internal kinetic energy at the microscopic level (cf. Leibniz's explana-

tion for the apparent loss of kinetic energy in inelastic impact mentioned in note 9). But the abstract point at issue did not demand acceptance of the kinetic theory, because the lost kinetic energy could alternatively be viewed as being stored in the 'tensions' of whatever force might be associated with heat.

15. Helmholtz used the word *Kraft*. This is now standardly translated as 'force' rather than 'energy,' but these two concepts were not clearly distinguished at the time, in either English or German. The general expectation at the time was that any conservation law would involve 'force' ('Kraft,' '*vis*'), where this was thought of as a directed quantity ('force of motion'), rather than as a scalarlike energy. (Here again we see the dominance of the Newtonian tradition, whose only conserved quantity was the vectorial momentum.) One of Helmholtz's most important contributions was to make it clear that even within the Newtonian tradition of rational mechanics it is the scalar energy that is conserved, rather than any vectorial 'force.' Even so, the confusions persisted for some time, as shown, for example, by Faraday's 1857 paper "On the Conservation of Force" (cf. Elkana [1974], pp. 130–8).
16. Indeed this line of thought seems to have become extremely popular in the late nineteenth century. The idea that the brain is a repository of 'nervous energy,' which gets channeled in various ways and is then released in action, is common among Victorian thinkers from Darwin to Freud.
17. My treatment of the conservation of energy raises some interesting questions in connection with quantum mechanics. (I am grateful to Barry Loewer for pressing these points on me.)

   (1) An initial query relates to my continued presentation of the issues in terms of forces. How does this fit in with modern quantum mechanics, which is normally formulated in terms of Hamiltonians rather than forces, that is, directly in energetic terms? But there is no substantial issue here, because the Hamiltonians can be seen as depending on the relevant forces (cf. McLaughlin [1992], p. 54).

   (2) On some interpretations, quantum systems do not always respect the conservation of energy. While energy is conserved in the 'Schrödinger evolution' of quantum systems, it is apparently violated by 'wave collapses.' Some, including myself, take this to argue against wave collapses. But, even if you don't go this way, it doesn't matter for this chapter, because (a) the argument from fundamental forces to completeness will still have weight even if conservation is restricted to Schrödinger evolutions, and (b) completeness is consistent with the indeterminacy of collapse outcomes, because the chances of those outcomes are still fixed by prior physical forces alone (cf. Note 2).

   (3) On some, but not all, collapse interpretations, however, *sui generis* factors do seem to fix whether a collapse occurs or not (even though the subsequent chances of the various possible outcomes then still depends entirely on the prior physical forces). I am thinking here of interpretations that say that collapses occur when physical systems interact with consciousness (or indeed that say that collapses occur when there are 'measurements,' or 'macroscopic interactions,' and then refuse to offer any physical reductions of these terms). On these interpretations, the completeness of physics is indeed violated, because collapses don't follow from more basic physical laws, but depend on 'emergent' causes. It would seem an odd victory for

nonphysicalists, however, if the sole locus of *sui generis* mental action were quantum-wave collapses.

18. However, not all emergentists were as sophisticated as Mill and Bain. In *Mind and its Place in Nature* (1923), C.D. Broad addresses the issue of whether independent mental causation would violate the conservation of energy (pp. 103–9). But instead of simply claiming that any mental force would operate conservatively, he insists that the principle of the conservation of energy does not explain all motions, even in physical systems, and so leaves room for other causes. He draws an analogy with a pendulum on a string, where he says that the "pull of the string" is a cause that operates independently of any flows of energy, and suggests that the mind might operate as a similar cause. While it is not entirely clear how Broad intends this analogy to be read, it is difficult to avoid the impression that he has mastered the letter of the principle of the conservation of energy, without grasping the wider physical theory in which it is embedded.
19. Indeed, and somewhat paradoxically, this species of 'bookkeeping' experiment may even have weighed in *favor* of postulating *sui generis* vital forces. This is because these experiments offer a counter to the argument from fundamental forces. That argument, remember, hinged on the claim that there is no direct inductive reason to suppose that any *sui generis* vital forces are conservative, if it is denied that they reduce to more fundamental forces. But experiments such as Rubner's do offer just such direct inductive reason, in that they show that any special forces operating inside bodies must always "pay back" just as much energy as they "borrow," even if they *don't* reduce to more fundamental forces. (I owe this point to Keith Hossack.)
20. McLaughlin ([1992], p. 89) attributes the end of British Emergentism, and therewith the rise of contemporary physicalism, to the 1920s quantum mechanical reduction of chemical forces to general physical forces between subatomic components. But it seems unlikely that this could have been decisive. After all, why should anybody be persuaded against special mental causes just because of the reduction of *chemistry* to physics? (Why should it matter to the existence of *sui generis* mental forces exactly how many independent forces there are at the level of atoms?) At most the reduction of chemistry to physics would have added weight to the argument from fundamental forces, by showing that yet another special force reduces to more basic forces. But it was irrelevant to the argument that I claim swayed thinkers in the twentieth century, the argument from physiology.
21. I would greatly like to thank Barry Loewer, Keith Hossack, and David Spurrett for comments on drafts of this chapter.