

Thinking about Consciousness

David Papineau

https://doi.org/10.1093/0199243824.001.0001

Published: 2002 **Online ISBN:** 9780191598166 **Print ISBN:** 9780199243822

END MATTER

Appendix: The History of the Completeness of Physics 3

Published: April 2002

Subject: Epistemology, Philosophy of Mind

Appendix:

A.1 Introduction

Those unsympathetic to contemporary materialism sometimes like to suggest that its rise to prominence since the middle of the twentieth century has been carried on a tide 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.

I reject this suggestion. In Chapter 1 I explained how materialism follows from a serious argument with persuasive premisses. Moreover, as I also intimated in Chapter 1, a proper appreciation of this argument indicates an alternative explanation of why philosophical materialism has only recently become so widespread. A crucial premiss in the causal argument is the \$\(\phi\) completeness of physics, and this premiss lacked convincing empirical support until well into the twentieth century. The reason why earlier philosophers were not materialists is not that they lacked some scientistic prejudice peculiar to the later twentieth century (after all, there were plenty of enthusiasts for science in previous centuries), but simply that they lacked the evidence which has now persuaded modern science of the completeness of physics.

In this Appendix I want to rehearse the history of scientific attitudes to the completeness of physics, and to show how changing views about this claim have interacted with attitudes to the mind-body problem. This will confirm my suggestion that modern materialism flows from the recent availability of the completeness of physics. But before I proceed, let me make one preliminary point. My claim that materialism derives from the completeness of physics might seem to be belied by the fact that few of the philosophers who developed modern materialism in the middle of the twentieth century, like the Australian central state materialists or David Lewis or Donald Davidson, made any explicit reference to this principle. If I am right that the completeness of physics was the crucial new factor, then we might have expected these philosophers to say so.

However, this is a relatively superficial worry. It is true that these founding fathers of modern materialism offered a number of variant arguments for materialism, and that not all of these arguments feature the completeness of physics as prominently as does the causal argument detailed in Chapter 1. Even so, it is not hard to see that nearly all these other arguments presuppose the completeness of physics in one way or

another, and would not stand up without it. The original defenders of materialism in the middle of the twentieth century may not have been explicit about the importance of the completeness of physics, but it remains the case that their innovatory views would not have been possible without it.

Now, these are all rather different arguments, and they give rise to rather different versions of materialism. But the point I want to make here is not sensitive to these differences. It is simply that none of these arguments would seem even slightly plausible without the completeness of physics. To see this, imagine that the completeness of physics were not true, and that some physical effects (the movement of matter in arms, perhaps, or the electrochemical changes which instigate those movements) were not determined by law by prior physical causes at all, but by *sui generis* non-physical 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 behaviour; (2) contra Armstrong and Lewis, it wouldn't necessarily be physical states which 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 behaviour are those connecting behaviour with physical antecedents, since there would also be laws connecting behaviour with mental antecedents.¹

A.2 Descartes and Leibniz

Let us now focus on the history of the completeness of physics. It may seem at first sight that the completeness of physics will follow from any version of physical theory which is formulated in terms of conservation laws. If the laws of mechanics tell us that important physical quantities are conserved whatever happens, then doesn't it follow that the later physical states of a system will always be 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.

p. 235 This is because Descartes's *quantity of motion* is a non-directional (scalar) 4 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 non-physical causes, and in particular *sui generis* mental causes, to alter the *direction* of a body's motion without violating Descartes's conservation principle. That 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, and then used this to explain how the mind could affect the brain without violating the conservation of 'quantity of motion' (Leibniz, 1898 [1696]: 327).

Now, there is little evidence that Descartes actually saw things this way, or 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.² And, in connection with the mind-body issue, they \Box persuaded him that there is no room whatsoever for mental activity to influence physical effects.³

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 '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 they collide. 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, in the sense of implying 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 that it is not necessary (cf. ch. 1 n. 2). So Leibniz's dynamics, unlike Descartes's, makes 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 1985). However, he did not infer mind-brain identity from his commitment to the completeness of physics. Instead, he adopted the doctrine of pre-established harmony, according to which the mental and physical realms are each causally closed, but pre-arranged by the divine will to march in step in such a way as to display the standard mind-brain correlations. In terms of the causal argument laid out in Chapter 1, Leibniz is denying the first premiss, about the causal influence of mind on matter. He avoids identifying mental causes with 4 physical causes, in the face of the completeness of physics, by denying that mental causes ever have physical effects.

A.3 Newtonian Physics

Some readers might now be wondering why this wasn't the end of the story. Given that Leibniz established, against Descartes, that both momentum and energy are conserved in systems of moving particles, why wasn't the history of the mind-brain argument already over? Of course, we mightn't nowadays want to follow Leibniz in opting for pre-established harmony, as opposed to simply embracing mind-brain identity. But this would simply be because we favour a different response to the causal argument laid out in Chapter 1, not because we have any substantial premisses that Leibniz lacked. In particular, the crucial second premiss 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 to hand in the second half of the seventeenth century, long before the rise of twentieth-century materialism?

Well, it was—but only on the assumption that 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 *non-mental* 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 was that of an *impressed force*. Rather than thinking of 'force' as something inside a body which might be transferred to other bodies in impact, as did all his contemporaries (and indeed most of his successors for at least a century⁴), Newton thought of forces as disembodied entities, acting on the affected body from outside. An impressed force 'a '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 1966 [1686]: 2, definition IV). Gravity was the paradigm. True, the force of gravity always arose from the presence of massive bodies, but it pervaded space, acting 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 non-physical mental forces among their causal antecedents. (Some readers may be feeling uneasy about the way in which the completeness of physics has now turned into an issue about what 'forces' exist. I shall address this issue at the end of this section.)

The switch from a pure impact-based mechanical philosophy to the more liberal world of Newtonian forces 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 impact. The Newtonian world of disembodied forces 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 \$\mathbb{L}\$ 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. 5

p. 239

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 below, 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 qeneris* mental forces.)

In any case, whatever the significance of later Newtonian derivations of the conservation of energy, early Newtonians themselves certainly saw no barrier to the postulation of *sui generis* mental forces. In a moment I shall give some examples. But first 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 \Box 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, since 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 *sui generis* 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 in Chapter 1, which defined the 'physical' as whatever can be identified without using specifically animate terminology—which then makes even deterministically governed *sui generis* mental forces come out 'non-physical', since they can't be so identified. This is the terminology which best fits with our original interest in the causal 4 argument for physicalism. We don't want deterministic mental forces to be counted as consistent with the 'completeness of physics', precisely because *this* kind of 'completeness of physics' wouldn't be any good for the causal argument: if mental forces are *part* of what makes 'physics' complete, then we won't be able to argue from this that mental forces must be identical with some *other* (inanimate) 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 a 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. ⁶

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 non-mental 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]: 211) 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 fibres.

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. Whereas 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 which acts according to strict laws. Ly Whytt can thus be seen as exemplifying my model of deterministic mental forces: the sentient principle

→ 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, in that its operations are fixed by a definite force law (Whytt 1755).

At this point let me say something about the terminology of 'forces' that I have been using in discussing Newtonian physics. It may be natural to present Newtonian physics in terms of reified forces in this way, but it is not mandatory. The alternative is to view the circumstances which supposedly generate these putative forces as themselves the direct causes of any resulting accelerations, and to regard the talk of 'forces' as simply a useful calculating device.

In an earlier paper about the history of the completeness of physics (Papineau 2000), I claimed that this choice made no difference to the issues, on the grounds that those who dispense with 'forces' can simply replace the question of whether there are 'mental forces' with the question of whether mental initial conditions ever make a difference to accelerations. (Cf. McLaughlin 1992: 64–5.) But now I think that the situation is more complicated, and that the reification of forces arguably makes it harder to uphold the completeness of physics.

The complication arises in connection with deterministic *mental* forces which are generated by special *physical* circumstances—for example, by circumstances found specifically within the brains of sentient beings. If we accept these as reified mental forces, then they would seem to violate the completeness of physics, since it seems that they will be needed as *sui generis mental* factors in any sufficient story about the causes of the accelerations they generate. On the other hand, if we refuse to reify forces, then a full story about the causes of those accelerations need mention only the prior *physical* circumstances which supposedly generate these 'mental forces', and the completeness of physics would thus seem to be respected. In such a case, then, the reification of forces seems to lead to a violation of the completeness of physics, where a non-reification does not.

The point is that while the non-reifiers may need special *laws* about the accelerations that are generated by the special physical circumstances found inside sentient bodies, laws that do not follow from other laws about accelerations, the antecedents of these special laws will still be physical, and so such antecedent causes will not violate the completeness of physics. By contrast, since those who reify forces do introduce *sui generis* mental forces to serve as causes of the relevant accelerations, their analysis of such accelerations will run counter to the completeness of physics. In short, special accelerations inside sentient brains would seem to violate the completeness of physics *if* we reify forces, but not otherwise.

p. 243 Fortunately, we can bypass this issue about the reification of forces here. This is because I shall be arguing that there are in any case no special accelerations inside brains, and so no reason, even for those who reify forces, to introduce special mental forces. So I shall be able to uphold the completeness of physics even on the assumption that forces should be reified. Such reification may make it harder to defend the completeness of physics. But, if there are in fact no special accelerations inside brains to motivate mental forces, then no such falsification will result.

Given this, I shall continue to talk in terms of 'forces' in what follows. Since this only makes my argumentative task harder, this will give me no unfair dialectical advantage. It simply sets me the greater challenge of showing that accelerations inside brains follow, not just from physical antecedents, but also in a way which is predictable from laws which also operate outside brains. (Or, as we would put it in terms of forces, that there are no forces inside brains which are not composed of physical forces which also operate outside brains).

A.4 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.

p. 244 A.4.1 Rational Mechanics

Through the eighteenth and early nineteenth centuries a number of mathematician-physicists, among whom the most important 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 to concentrate on the effects produced by other forces. (See Elkana 1974: ch. 2 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 $1/2\Sigma mv^2$) 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 no 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 himself 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 holds only under the assumption that all forces are conservative. We nowadays take this requirement to be satisfied for all fundamental forces. But again, this was no part of eighteenth-century thinking. Some familiar forces happen to be conservative, but plenty of other forces are not. Gravitation, for example, is conservative, since it depends only on the positions of the particles, and not on their velocities, or the elapsed time. But, by contrast, frictional forces are not conservative, since 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.

p. 245 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 which falls out of the equations when the operative forces all happen to fall within a subset of possible forces. (Cf. Elkana 1974: ch 2.)

A.4.2 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, say), 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 non-conservative forces like friction and other dissipative forces need not be non-conservative after all, since the kinetic energy apparently lost when they act will in fact be preserved by the heat energy gained by the resisting medium.⁹

The stage was now set for the formulation of a universal principle of the conversation of energy. We can distinguish three elements which 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 which can manifest itself in different forms. Third, these experiments also suggested that apparently non-conservative forces like 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 4 connections. In 1847, at the age of 26, he published his monograph *Über 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', again showing that there are fixed equivalences between these 'forces', heat, and mechanical energy.¹⁰

A.4.3 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 programme in physiology, aiming to show that phenomena like respiration, animal heat, and locomotion could all be understood to be governed by the same laws as operate in the inorganic realm.

It seems likely that it was Helmholtz's specific combination of physiological interests and sophisticated physical understanding that precipitated his crucial synthesis of the different strands of research feeding into the conservation of energy. His desire to bring living systems under a unified science allowed him to see 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. ¹¹

A.4.4 Vital Forces

Helmholtz was part of a tradition in experimental physiology which set itself in opposition to the previous generation of German *Naturphilosophen*. During the eighteenth century the Newtonian categories of 'irritability' and 'sensibility' had gone through various transformations, and by the end of the century were widely referred to under the heading of *Lebenskraft*, or 'vital force', though there was continued disagreement on the precise nature of such forces. Meanwhile, within the tradition of German idealism, the notion of vital force had broken loose from its original Newtonian moorings, and became part of a florid metaphysics imbued with romanticism and idealism.

According to the *Naturphilosophen*, organic matter was infused with a special power which organized and directed it. Following Blumenbach and Kant, Schelling took up the term *Bildungstrieb* ('formative drive'), because of the excessively mechanical connotations he discerned in the traditional term *Lebenskraft*. Schelling and the other *Naturphilosophen* viewed this formative drive 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 towards that ideal form. (See Coleman 1971: ch. 3; Steigerwald 1998.)

The experimental tradition which included Helmholtz can be seen as a reaction to these extravagant 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 might 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 they were reverting to the tradition of eighteenth-century physiology. They viewed vital forces simply as special Newtonian forces, additional to gravitational forces, chemical forces and so on, 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. 6; Elkana 1974: ch. 4.)

A.4.5 Does the Conservation of Energy Rule out Vital (and Mental) Forces?

The interesting question, from our point of view 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 himself, and his young colleagues from Müller's laboratory, were committed to the view that no forces operated inside living bodies that are not also found in simpler physical and chemical contexts (Coleman 1971: 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 I earlier distinguished two ways in which early Newtonian theory might allow room for such *sui generis* animate forces. 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, say.) Why shouldn't such a force simply respect the conservation of energy by not causing accelerations which 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 buildups of potential energy, and vice versa. But we couldn't really speak of a 'buildup' 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 which 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 which *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 exactly which basic forces exist. It requires only that, whatever they are, they operate deterministically and conservatively.¹²

A.5 Conservative Animism

In this section I shall briefly sketch the evolution of attitudes to the completeness of physics since Helmholtz's promulgation of the universal conservation of energy. The issues are not straightforward, and there is no question of dealing with them fully here. But I would like to offer at least an outline of how the argument for the completeness of physics has developed since the mid-nineteenth century.

Helmholtz's doctrine left various options open in relation to the completeness of physics. For a start, you could simply deny that the conservation of energy applied to animate forces. That is, you could hold that vital and mental forces are an exception to the general rule that all forces are conservative, and thus insist that the conservation of energy holds only when we are dealing with inanimate forces.

However, this option does not seem to have been popular among \$\(\) scientifically informed commentators in the second half of the nineteenth century. The doctrine of the universal conservation of energy won widespread acceptance within a decade or two of its formulation. There is of course an evidential question here too: how far was this almost immediate agreement on the conservation of energy dictated by the strength of evidence rather than by intellectual fashion? But there is no question of pursuing this issue here. So let me assume for present purposes that the conservation of energy itself was well supported by the middle of the nineteenth century, and focus instead on where this left the completeness of physics.

Certainly this is how the writers I shall discuss henceforth saw the matter. Their question was not whether energy is always conserved, but rather, whether such conservation leaves any room for animate forces.

As I pointed out in the last section, it is clear that conservation does leave such room. The universal conservation of energy may rule out indeterministic animate forces, but there is clearly nothing in it to preclude deterministic animate forces that do respect the conservation of energy. Even so, as I observed, Helmholtz and his young colleagues rejected any such special animate forces. It is interesting to consider what might have persuaded them of this. I suspect that they were moved by what I shall call 'the argument from fundamental forces'. This is the argument that all apparently special forces characteristically *reduce* to a small stock of basic physical forces which conserve energy. Causes of macroscopic accelerations standardly turn out to be composed of a few fundamental physical forces which 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, like all causes of physical effects, are ultimately composed of the few basic physical forces.

It is possible that this line of thought was influential in originally persuading Helmholtz of the universal validity of the conservation of energy. We have already seen how Helmholtz's initial formulation of this principle hinged on the assumption that friction and other dissipative forces are non-fundamental forces, macroscopic manifestations of processes involving more fundamental conservative forces. For it is only if we see macroscopic forces like friction as reducing to fundamental conservative forces that we can uphold the universal conservation of energy. Given this view about dissipative forces, a natural move would be to generalize inductively and conclude that *all* apparently special forces must reduce to a small stock of fundamental forces. After all, those special forces which have been quantitatively analysed, like friction, p. 251 turn out to reduce to more fundamental conservative forces. So this could be seen as \$\(\) providing some inductive reason to conclude that any other apparently special forces, like muscular forces or vital forces or mental forces, will similarly reduce.

Thus consider how Helmholtz argues in *Über die Erhaltung der Kraft*. He takes pains to stress how it is specifically *central* forces independent of time and velocity which ensure the conservation of energy. This emphasis on central forces (by which Helmholtz meant forces which act along the line between the interacting particles) now seems dated. Nowadays conservativeness is normally defined circularly, as a property of those forces which do no work round a closed orbit. This definition does not require a restriction to central forces. However, Helmholtz was in no position to adopt the circular modern definition of conservativeness. He was aiming to *persuade* his readers of the general conservation of energy, so needed an argument. It wouldn't have served simply to observe that energy is conserved by those forces which 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 that energy is conserved by all forces, *unless* all forces are central. Why should this be? Well, as above, one persuasive thought would be that there is a small stock of basic central forces, and that all causes apparently peculiar to special circumstances are composed out of these.

It is clear from our earlier discussion, however, that this reductionist move is not *essential* to a commitment to the universal conservation of energy. An alternative strategy would be to allow that there are *sui generis* animate forces, and to maintain that these fundamental special forces are conservative in their own right. True, this position is open to the objection that there is no direct reason to suppose that any such *sui generis* animate forces *will* be conservative, if they do not reduce to other fundamental conservative forces. But this could be countered with the alternative inductive thought that, since all the *other* fundamental forces so far examined have turned out to be conservative, we should infer that any extra vital or mental fundamental forces will be conservative too.

Somewhat oddly, physiological research in the second half of the nineteenth century added support to this anti-reductionist stance, by offering direct empirical evidence that if there were any special animate forces, they would have to respect the conservation of energy. In a moment I shall argue that physiological research has also given us strong reason to doubt that there *are* any special animate forces. But this latter conclusion derives from investigations at a microscopic cellular level, and such research had to wait until the twentieth century. Prior to that, however, there was a flourishing tradition of energetic research at a more was a macroscopic level, which identified chemical and energetic inputs and outputs to various parts of the body, and showed that animals are subject to general conservation principles. Especially noteworthy were Max Rubner's elaborate 1889 respiration calorimeter experiments, which showed that the energy emitted by a small dog corresponds exactly to that of the food it consumes. (See Coleman 1971: esp. 140–3.)

The interesting point is that this kind of research did nothing to support the reductionist view that all apparently special forces reduce to a few basic inanimate 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 animate creatures, but operate in such a way as to 'pay back' all the energy they 'borrow', and vice versa.

Rather, research like Rubner's would have added weight to the position of those who took the existence of sui generis animate forces to be consistent with the conservation of energy, as further items in the category of fundamental conservative forces.

As exemplars of this position, 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. And 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. 13 (The 'British Emergentists' discussed by McLaughlin constituted a philosophical movement committed precisely to non-physical causes of motion in my sense, causes which were not the vectorial 'resultants' of basic physical forces like gravity and impact, but which 'emerged' when matter arranged itself in special ways. The particular idea which attracted Mill and Bain was that these 'emergent forces' might be stored as unrealized potentials, ready to manifest themselves as causes of motion only when the relevant special circumstances arose. 14)

A.6 The Death of Emergentism

McLaughlin explains how British Emergentism continued to flourish well into the twentieth century. This highlights the question at issue in this Appendix. 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 disfavour?

Here I think we need to refer to a second line of argument against such forces, an 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 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 which 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.

As I intimated above, the relevant research dates mostly from the twentieth century. Physiological research in the nineteenth century did not penetrate to the level of forces operating inside bodies. However, 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 reductionist argument from fundamental forces, 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 analysed, culminating in the discovery of DNA. In the same period, neurophysiological research mapped the body's neuronal network and analysed the electrical mechanisms responsible for neuronal activity. Together, these developments p. 254 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 the completeness of physics against the background provided by the argument from fundamental forces. One virtue of this

explanation in terms of these 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. Imagine a ranking of different thinkers through this period, in terms of the amount of physiological evidence 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 the twentieth 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 buildup of support for the completeness of physics as the physiological evidence accumulated, culminating, I would contend, in a general scientific consensus by the 1950s.

Brian McLaughlin offers a rather different explanation for the demise of British Emergentism. He attributes it to the 1920s quantum-mechanical reduction of chemical forces to general physical forces acting on subatomic components (1992: 89). But it seems unlikely that this could have been decisive. After all, why should anybody who was still attracted to *sui generis* animate forces in the 1920s have turned against them simply because of the reduction of *chemistry* to physics? Why should it have mattered to them exactly how many independent forces there were at the level of atoms? At most, the reduction of chemistry to physics would have added some marginal weight to the argument from fundamental forces, by showing that yet another special force reduces to more basic forces. But anybody who had resisted the argument from fundamental forces so far, still upholding vital and mental forces as extra members of the pantheon of fundamental forces into the twentieth century, would surely not be bowled over simply because the physical theorists had now modified the precise inventory of forces operating at the atomic level. To understand why British Emergentism lost ground over the first half of the twentieth century, we need to recognize a different kind of argument: namely, the argument from the emerging findings of physiological research. ¹⁵

A.7 Conclusion

This Appendix has charted the history of changing attitudes to the completeness of physics. The important point is that a scientific consensus on completeness was reached only in the middle of the twentieth century. In earlier centuries there was no compelling reason to believe that all physical effects are due to physical causes, and few scientists did believe this. But by the 1950s the combination of the physiological evidence with the argument from fundamental forces left little room for doubt about the doctrine.

In Chapter 1 I raised the question of why philosophical physicalism is peculiarly a creature of the late twentieth century. I hope I have now succeeded in showing that this is no intellectual fad, but a reflection of developments in empirical theory. 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 premiss could be slotted into a number of variant arguments for physicalism. There seems no 4 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 which initially created a scientific consensus on completeness in the 1950s.

Perhaps this is what Tyler Burge has in mind when he says that 'materialism is not established, or even deeply supported, by science', or Stephen Clark when he doubts whether anyone could 'rationally suppose' that empirical evidence 'disproves' mind-body dualism. If so, there is no more I can do to persuade them of the completeness of physics. However, I see no virtue in philosophers refusing to accept a premiss which, by any normal inductive standards, has been fully established by over a century of empirical research.

Notes

- In other writings from the middle of the century, the relevance of the completeness of physics does not need to be excavated, since it lies on the surface. Thus see Feigl 1958, Oppenheim and Putnam 1958.
- 2 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.')
- I am here using 'physical' in the sense it would have been understood by seventeenth-century mechanical philosophers, as referring to primary properties like mass and motion, and to anything ontologically determined thereby. In the next section I shall consider the Newtonian system, which allows an open-ended range of primary properties; in that context I shall use 'physical' to refer more generally to those properties which are 'inanimate' in the sense of Chapter 1.
- 4 Cf. Papineau 1977.
- 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 which 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 themselves do not rule it out. (It is also worth noting that perpetual motion was no by means universally rejected by seventeenth- and eighteenth-century physicists. Cf. Elkana 1974: 28–30.)
- 6 Here I am closely following Steigerwald 1998: ch. 2.
- Moreover, it is probably the right way to talk anyway. While early Newtonian physics can avoid reifying forces, modern field theories cannot.
- Note that those who don't reify forces will be able to run a completeness-based causal argument for mind-brain identity even if cerebral accelerations are governed by special laws independent of other laws about accelerations. In this connection, it is useful to distinguish between 'weak' and 'strong' reduction. A weak reduction requires only that mental causes be identified with physical causes. A strong reduction requires also that the laws by which such causes operate follow by composition from non-special laws. Those who admit special cerebral accelerations but don't reify forces will only be able to achieve a weak reduction. If we do reify forces, on the other hand, then there is no room for a reduction which is weaker than a strong reduction: for once we introduce forces, there will be mental causes (forces) which are distinct from physical causes as soon as there are special mental laws which do not follow by composition from non-special laws. As I said, I myself have no need to make space for weak reduction by avoiding forces, since I shall argue that there are no special mental laws, and thus that strong reduction can be upheld anyway. (I think I remember May Brodbeck making this distinction between weak and strong reduction in the 1960s; but I haven't been able to track down a reference.)
- 9 One model for this preservation was the kinetic theory of heat, which took the macroscopic kinetic energy apparently lost to be converted into internal kinetic energy at the microscopic level (cf. Leibniz's explanation for the apparent loss of kinetic energy in inelastic impact mentioned in footnote 2 above). But the abstract point at issue did not demand

- acceptance of the kinetic theory, since the lost kinetic energy could alternatively be viewed as being stored in the 'tensions' of whatever force might be associated with heat.
- 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, neither in English nor 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'=momentum), rather than as a scalar like energy. (Here again we see the dominance of the Newtonian tradition, in which the 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: 130–8.)
- In Papineau (2000) I suggested that their lack of physiological interests meant that the experimenters investigating mechanical equivalences, like Joule, had no interest in articulating a universal principle governing all natural interactions. However, this is wrong about Joule, who did defend such a universal principle. (Cf. Smith 1998: ch. 4.)
- 12 I have the impression that scientifically informed late nineteenth-century philosophers were not particularly exercised by our issue of whether or not there are special vital or mental forces. Understandably enough, they were far more interested in the determinism implied by the conservation of energy even on the assumption of special animate forces. Cf. Tyndall 1898 [1877].
- 13 Indeed this line of thought seems to have become extremely popular in the late nineteenth century. The idea that the brain in a repository of 'nervous energy', which gets channelled in various ways, and is then released in action, is a commonplace of Victorian thinkers from Darwin to Freud.
- 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 which operates independently of any flows of energy, and he 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 conservation of energy, without grasping the wider physical theory in which it is embedded.
- I have presented the conservation of energy as a boundary principle which limits the range of possible forces and helped rule out special animate forces. Barry Loewer has pressed me on whether this emphasis on the conservation of energy is consistent with modern quantum mechanics. Let me make two brief comments. (1) 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, the conservation of energy will still be respected in Schrödinger evolutions, and this will itself limit the range of (non-collapse-causing) forces. (2) On some, but not all, collapse interpretations, distinctive special causes will be responsible for whether a collapse occurs or not (even though the subsequent chances of the various possible outcomes will still depend entirely on prior physical forces). I am thinking here of interpretations which say that collapses occur when physical systems interact with consciousness (or indeed which say that collapses occur when there are 'measurements', or 'macroscopic interactions', and then refuse to offer any physical explanations of these terms). On these interpretations, the completeness of physics will be violated, as well as the conservation of energy, since collapses don't follow from more basic physical laws, but depend on 'emergent' causes. It would seem an odd victory for anti-materialists, however, if the sole locus of sui generis mental action were quantum wave collapses.