CHAPTER 2

THE CAUSAL CLOSURE OF THE PHYSICAL AND NATURALISM

DAVID PAPINEAU

Introduction

OVER the latter half of the last century English-speaking philosophy became increasingly committed to naturalistic doctrines. Much of this naturalistic turn can be attributed to the widespread acceptance of the thesis that the physical realm is causally closed.

This chapter will contain four sections. Section 2.1 offers an initial formulation of the thesis that physics is causally closed and explains its immediate implications. Section 2.2 then discusses the evidence for the thesis from a historical perspective. Section 2.3 considers ways of making the thesis properly precise. Finally, Section 2.4 explores the connections between the thesis and the more general issue of naturalism.

2.1 Causal Closure and its Immediate Implications

At first pass the causal closure of physics says that every physical effect has a sufficient physical cause. If this thesis is true, it distinguishes physics from all other subject

domains. The biological realm, for example, is not causally closed in this sense, since biological effects often have non-biological causes, as when the impact of a meteorite precipitated the extinction of the dinosaurs. Again, meteorology is not causally closed: the burning of carbon fuels—a non-meteorological event—is causing global warming. Nor, importantly, is the mental realm causally closed: a mental pain can be caused by sitting on a physical drawing pin, or a train of thought can be interrupted by a loud noise.

Physics, by contrast, does seem to be causally closed. If you consider any physical effect, then there will arguably always be some prior sufficient physical cause: we expect to be able to account for physical effects without leaving the physical realm itself. In particular, this seems to hold even for physical effects which take place within the bodies of conscious beings. When the muscle fibres in my arm contract, this is presumably due to electrochemical activity in my nerves, which is due to prior physical activity in my motor cortex, and so on. In principle, it would seem possible to account for this entire sequence solely in terms of the resources offered by physics itself, and without making any essential appeal to any other subject matter. (Isn't the causal closure of physics undermined by quantum-mechanical indeterminism? Well, quantum mechanics does call for a more careful formulation of the causal-closure thesis, but this doesn't matter for the philosophical implications. Let us shelve this complication until Section 2.3.)

The causal closure of physics is solely a claim about how things go within physics itself. It does not assert that everything is physical, but only that those things that *are* physical have a physical cause. So it does not rule out realms of reality that are quite distinct from the physical realm. It is entirely consistent with the causal closure of physics that there should be self-sufficient realms that operate quite independently of physical goings-on (a realm of ghostly spirits, say). The causal closure of the physical says only that when we are within the physical realm we will find that every physical effect has a physical cause.

Even so, the causal closure of the physical does give rise to a powerful argument for reducing many prima facie non-physical realms to physics: for it indicates that *anything that has a causal impact on the physical realm* must itself be physical. This is because the causal closure of the physical seems to leave no room for anything non-physical to make a causal difference to the physical realm, since it specifies that every physical effect already has a physical cause.

To see the significance of this point, note that many prima facie non-physical realms—such as the biological, meteorological, and mental realms—certainly do seem to have physical effects. Not only are these realms causally affected by physical events, as noted above, but they in turn exert a causal influence on the physical realm. An infection (biological) can cause a physical rise in temperature; a hurricane (meteorological) can destroy physical houses; my current thoughts (mental) can give rise to physical movements of my fingers on a keyboard. However, the causal closure of the physical says that these physical effects must already have physical causes: the rise in temperature will be caused by cellular-level chemical processes; the houses' destruction by fast-moving gases; the movements of my fingers by neuronal activity

in my brain. How then can the infection, the hurricane, and my thoughts also make a causal difference? The causal closure of the physical would seem to leave no room for these prima facie non-physical causes to matter to the physical effects. Unless, that is, we identify the prima facie non-physical causes with the physical causes. If we equate the infection, hurricane, and thoughts with the cell-level activity, fast-moving gases, and neuronal activity respectively, then their causal efficacy will be respected. The biological, meteorological, and mental causes won't be eclipsed by the physical causes, simply because they will be one and the same as the physical causes—they will be 'non-physical' only in the sense that they are normally referred to using specialist (biological, meteorological, mental) terminology, and not because they are ontologically different.

The thesis of the causal closure of the physical thus argues that many prima facie non-physical occurrences—all those that exert an influence on the physical realm—must themselves in fact be physical. For otherwise it is hard to see how they could have any physical effects.

2.2 A Historical Perspective on the Evidence for the Causal Closure of Physics

Why should we believe in the causal closure of physics (which for the moment I shall regard as the simple claim that every physical effect has a sufficient physical cause)? It doesn't look as if this is an a priori matter—there is nothing conceptually contradictory in the idea that physical phenomena may be affected by non-physical causes, as Descartes supposed, for example. So the causal closure of physics, if true, must somehow follow from the findings of science. But exactly which findings? Which part of science, if any, argues that the physical is causally closed?

At first sight it might seem as if causal closure follows from the presence of conservation laws in physics: if there are laws specifying that important physical quantities stay constant over time, won't this show that the later values of physical quantities must be determined by earlier values? However, it depends on what conservation laws you have. Not any set of physical-conservation laws rules out non-physical causes for physical effects.

Consider Descartes's early seventeenth-century physics. This was based on a conservation law, but nevertheless allowed the non-physical mind to affect the physical brain. The central principle of Descartes's physics was the conservation of *amount of motion*, which Descartes took to be the product of the masses of all bodies and their scalar *speeds*. (So amount of motion is different from momentum, which is the product of mass and vectorial *velocity*: a car going round a bend at a constant speed conserves amount of motion but not momentum.) In line with Descartes's mind–body interactionism, the conservation of amount of motion leaves plenty of

room for non-physical causes to intrude on the physical realm. In particular, if mental causes (operating in the pineal gland?) cause particles of matter to change their direction (but not their speed), then this will not in any way violate the conservation of amount of motion.

Descartes's physics might allow an independent mind to affect the brain, but Descartes's physics is wrong. Later in the seventeenth century Leibniz replaced Descartes's law of the conservation of 'motion' with the two modern laws of conservation of (vectorial) momentum and of (scalar) kinetic energy, and thereby arrived at what we now regard as the correct laws governing impacts. Now, Leibniz's physics, unlike Descartes's, did indeed imply that the later values of *all* physical quantities are determined by their earlier values, and therewith the causal closure of the physical. Leibniz thus faced the argument outlined in my first section above: How can a non-physical mind have physical effects, given that the causal closure of the physical seems to leave no room for anything non-physical to make a causal difference to the physical realm (see Woolhouse 1985)? However, Leibniz did not draw the modern physicalist conclusion that the mind must therefore be identical to the brain. Rather, he denied that the mind in fact has any physical effects. Since it seemed incontrovertible to him that mind and brain must be ontologically separate, he instead inferred from the causal closure of the physical that the mind is impotent to affect the physical world. (It only appears to do so because of the 'pre-established harmony' with which God has arranged both the mental and physical worlds.)

While Leibniz's physics implies the causal closure of the physical, this is not true of the Newtonian system of physics that replaced it at the end of the seventeenth century. The crucial difference is that where Leibniz upheld the central principle of the 'mechanical philosophy' and maintained that all changes of velocity are due to impacts between material particles, Newton allowed that accelerations can also be caused by disembodied forces, such as the force of gravity. Moreover, Newton's system was quite open-ended about the range of different forces that exist. In addition to gravity, Newton and his followers came to recognize magnetic forces, chemical forces, forces of adhesion—and indeed vital and mental forces, which arose specifically in living bodies and sentient brains respectively. If we count vital and mental forces as 'non-physical' (we shall return to this point in the next section), then the admission of such forces undermines the causal closure of the physical. For it means that physical effects, in the form of accelerations of particles of matter, will sometimes be due to the operation of non-physical vital or mental causes. Given this, Newtonian science once more allowed that a non-physical mind can influence the physical world.

Doesn't Newtonian physics coincide with Leibniz's at least to the extent of upholding the conservation of momentum and energy? So why don't these principles imply the causal closure of the physical, as they did for Leibniz? The answer is that Newton's physics differs from Leibniz's in the way the conservation of energy must be understood. Leibniz upheld the conservation of *kinetic* energy. However, the existence of Newtonian force fields means that this quantity is not always conserved: for example, two bodies receding from each other will slow down due to their mutual gravitational

attraction, and so lose kinetic energy. Newtonian energy conservation specifies rather that the sum of kinetic *and* potential energy is constant. Potential energy is the latent energy stored when bodies are 'in tension' in force fields, as when two receding gravitating bodies cease to move apart and are about to accelerate together again. The notion of potential energy was not prominent in early Newtonian physics, but by the middle of the nineteenth century physicists concluded that all forces operated so as to conserve the sum of potential and kinetic energy—any loss of kinetic energy would mean a rise in potential energy, and vice versa.

The emergence of this modern version of the 'conservation of energy' placed strong restrictions on what kinds of forces can exist, but it by no means rules out vital and mental forces. Provided that the fields of these forces store in latent form any losses of kinetic energy they occasion (cf. the notion of 'nervous energy'), the presence of non-physical forces will be perfectly consistent with the conservation of kinetic plus potential energy. True, the conservation of kinetic plus potential energy does apparently imply that all forces must be governed by deterministic force laws (otherwise what would ensure that they always 'paid back' any kinetic energy they 'borrowed'?), and this greatly exercised many Victorian thinkers, especially given that nothing in early Newtonian physics had ruled out spontaneously arising mental forces. But, even so, the Newtonian conservation of energy does not stop *deterministic* vital and mental forces affecting the physical realm.

Nevertheless, during the late nineteenth and the twentieth centuries an increasing number of scientists have come to doubt the existence of vital and mental forces. The most significant evidence seems to have come directly from physiology and molecular biology, rather than from physics. Over the last hundred and fifty years a great deal has come to be known about the workings of biological systems (including brains), and there has been no indication that anything other than basic physical forces is needed to account for their operation. In particular, the twentieth century has seen an explosion of knowledge about processes occurring within cells, and here too there is no evidence of anything other than familiar physical chemistry. The result has been that the overwhelming majority of scientists now reject vital and mental forces, and accept the causal closure of the physical realm (see Papineau 2002: app.; see also McLaughlin 1992).

2.3 The Causal-Closure Thesis Refined

The causal-closure thesis presupposes some prior concept of the *physical* realm. Some commentators argue that the unclarity of this concept empties the causal-closure thesis of content (Crane and Mellor 1990).

At first sight it might seem that 'physical' can be defined by reference to the fundamental categories of physical *theory*. However, this strategy generates an awkward dilemma for advocates of the causal-closure thesis (Hempel 1980). On the one hand, they can equate 'physical' with the category of phenomena recognized by *current*

physical theory. But then it seems implausible that 'physics' in this sense is closed: past form suggests very strongly that physics will in time come to posit various new fundamental causal categories. Alternatively, advocates of causal closure might wish to equate 'physical' with the ontology of some ideal *future* physical theory. But then it is hard to see how the causal closure of the 'physical' could have any current philosophical significance, given that we are as yet quite ignorant of exactly what this ideal future physical theory will include.

However, this dilemma is by no means inescapable. True, neither current physics nor ideal future physics gives us a suitable notion of 'physics' for framing the causal-closure thesis. But this does not mean there are no other notions of 'physics' that will serve this purpose. Indeed there are arguably a number of *different* ways of understanding 'physics' that will yield a well-evidenced and contentful causal-closure thesis (Jackson 1998: 7–8).

For a start, we could simply define 'physical' as *not sui generis mental or biological*. This understanding of 'physical' was in effect assumed at the end of the last section, in the argument that the non-existence of vital or mental forces establishes the causal closure of physics. Note that nothing in that argument assumed a definitive list of fundamental physical categories; rather the thought was simply that this list would *not* include any *sui generis* mental or vital entities. This is a relatively inclusive understanding of 'physical': it counts as a 'physical' cause anything that isn't mental or vital, and to this extent renders the causal closure of the physical a relatively weak thesis. But even so it remains a thesis of much philosophical interest, since it still argues that any mental or vital causes of physical effects must be identical to causes that are not themselves *sui generis* mental or vital.

A rather stronger reading of 'physical' would take it to cover any categories of the *same general kind as are recognized by current physical theory*. The idea here is to understand 'physical' as referring to some mathematically characterizable set of primary qualities, recognizably descended from the categories of modern physics, even if not identical to them. On this reading the list of fundamental 'physical' categories will be taken to include not just anything non-vital-or-mental, but more specifically items that display the same kind of simplicity as the categories of contemporary physics. Again, there seems good reason to suppose that 'physics' in this sense is casually closed, and therefore that anything which in this sense has 'physical' effects must itself be 'physical'.

Finally, and even more specifically, there is the option of equating 'physical' with *microscopic*. Modern physical theory characteristically deals with fundamental entities that are very small. Given this, we might define 'physical' as referring to anything that is composed of entities below a certain size (Pettit 1993). The causal-closure thesis will then hold that every effect that is microscopically constituted must have a cause that is similarly microscopically constituted. Again, this thesis seems plausible, and argues that everything that has microscopically constituted effects must itself be microscopically constituted.

Apart from worries about the meaning of 'physical', there are various other complications with the causal-closure thesis. So far this thesis has been presented fairly informally. A more careful formulation would be this:

> Every physical effect has an immediate sufficient physical cause, in so far as it has a sufficient physical cause at all.

Consider first the requirement that the physical cause be 'immediate'. This is needed to rule out the possibility of physical causes which produce their physical effects only via non-physical intermediaries. Imagine, for example, that every bodily movement is indeed fully determined by some prior brain state, but that these brain states only produce the bodily movements *indirectly*, by first producing some *sui generis* mental effect, which then affects the nervous system, and so on (see Lowe 2000). This set-up would ensure that every bodily movement had some (indirect) sufficient physical cause—namely, the prior brain state—but intuitively it wouldn't ensure the causal closure of physics, since it wouldn't require that the direct causes of physical effects must be physical, and therefore wouldn't ensure that every physical effect had a *purely* physical prior history. The requirement that every physical effect have an *immediate* physical cause will rule out any mixed physical/mental/physical histories of this kind, since it ensures that any physical effect has an immediately prior physical cause, which in turn has an immediately prior physical cause, and so on.

Now consider the requirement that the physical cause be 'sufficient'. This is needed to ensure that it causes the physical effect by itself, and not solely in virtue of its *conjunction* with some *sui generis* non-physical cause. Imagine, for example, that some neuronal activity is caused by the conjunction of some chemical state *and* some *sui generis* mental cause. Then that neuronal activity would have *a* physical cause (the chemical state), but this cause would not have sufficed on its own, in the absence of the *sui generis* mental factor. This is clearly less than we need for a philosophically significant closure thesis. To make sure we have the right kind of closure thesis, we thus need to require that every physical effect have a physical cause that suffices on its own.

However, this last requirement now highlights the need for the final qualification entered above—that every physical effect has a sufficient immediate physical cause 'in so far as it has a sufficient immediate cause at all'. The reason for this latter qualification is to accommodate the indeterminism of modern quantum mechanics, which tells us that certain physical effects are random, without any sufficient determining cause.

At first sight it might seem as if this qualification will empty the causal-closure thesis of any interest, and in particular prevent it from implying that everything with physical effects must itself be physical. For it might seem that quantum indeterminism creates room for *sui generis* non-physical causes (operations of the will, perhaps) to exert a 'downwards' influence on the physical realm, by influencing whether or not random quantum events occur. However, this appearance is illusory. Quantum mechanics still specifies that random physical effects have their *probabilities* fixed by sufficient immediate physical causes. If we understand the causal closure of the physical as covering this kind of physical determination of physical probabilities,

then it will once more rule out any *sui generis* non-physical cause for a physical effect. For any such *sui generis* cause would have to make a difference to the *probability* of the relevant physical effect, and this would once more run counter to the causal-closure thesis, understood now as implying that this probability is already fixed by some sufficient immediate physical cause.

2.4 Causal Closure and Naturalism

At its most general, philosophical naturalism is the view that philosophy is continuous with the natural sciences. This view can usefully be divided into two parts: *methodological* naturalism, which asserts that philosophy uses the same methods of investigation as the natural sciences, and *ontological* naturalism, which says that the subject matter of philosophy coincides with that of the natural sciences. The thesis that physics is causally closed bears on both these aspects of naturalism. In this section I shall first comment briefly on the relation between causal closure and methodological naturalism; after that I shall turn to its connection with ontological naturalism.

The most immediate way in which causal closure bears on methodological naturalism is by itself providing an important illustration of the methodological thesis. Recall how Section 2.2 showed the causal-closure thesis to be a highly empirical claim, whose acceptance derives from detailed empirical evidence about the causes of physical effects. This clearly does not stop the causal-closure thesis from having substantial philosophical consequences for the nature of mind and other philosophically interesting categories. Methodological naturalists will say that this is a prime example of the way that philosophical results, just like scientific results, characteristically depend on empirical findings.

The causal closure of physics also promises to bring philosophy into close contact with science at a more detailed level. The causal-closure thesis promises to show that the general run of prima facie non-physical entities are in fact physical. If this is right, then we can expect specific scientific findings to illuminate many familiar subjects of philosophical enquiry. To the extent that philosophically interesting categories are physically constituted, we can look to scientific investigation to uncover their underlying physical natures.

Exactly how far scientific investigation will be able illuminate such categories, however, will depend on whether the causal-closure thesis really does show them to be physically constituted, and if so in what precise sense. So far we have not looked very closely at the argument from causal closure to physical constitution. There is plenty of room for debate about the range of items to which this argument applies, and about the precise sense in which it shows these items to be physical. This now brings us to the bearing of the causal-closure thesis on *ontological* naturalism. In the rest of this final section I shall examine exactly what the causal-closure thesis implies about the physical constitution of reality.

It will be convenient to proceed in two stages. First I shall consider those items to which the causal-closure thesis applies directly; namely, those entities that have physical effects. After that I shall consider whether the causal-closure thesis has any implications for entities that are not themselves physically efficacious.

Take some prima facie non-physical cause that has a physical effect. According to the causal-closure thesis, this physical effect already has a sufficient physical cause. So, on pain of deeming this effect to have two independent causes, we need somehow to collapse the non-physical cause into the physical cause. But what exactly does this require? There are different views about what is needed to render the two causes indistinct.

A minimum requirement is that the two causes be *token identical*, in the sense that they both involve the same spatio-temporal particulars. If some thought and some brain state are both to be counted as the cause of my fingers moving over the keyboard, say, then it seems clear that these causes cannot involve spatio-temporally divergent particulars. If they did, then we would have a *substantial* dualism of causes, a pair of causes involving different particular substances.

Still, it is generally agreed that token identity on its own is not enough to rule out a divergence of causes. It may rule out 'substance dualism', but it does not yet exclude 'property dualism'. To see why, note that token identity is a very weak doctrine, since it does not imply any relationship at all between the *properties* involved in the physical and non-physical causes—it is enough that the same particular entity should possess both these properties. In this sense, my height is token identical with my weight, since these properties are both possessed by the same particular; namely, me. Yet it seems wrong to conclude on this account that my height causes what my weight causes. Similarly, it seems unwarranted to conclude that my thoughts cause what my brain states cause, simply on the grounds that these causes must involve the same particulars. Causes are in some sense property-involving, and any claim that two causes coincide will require that the properties they involve are in some sense also coincident.

So much is widely agreed by contemporary philosophers of mind. At this point, however, consensus ends. One school holds that the causal argument requires *type identity*, the strict identity of the relevant properties of the non-physical cause with some physical properties of the physical cause. They are opposed by those who say that it is enough if the non-physical properties *metaphysically supervene* on the physical ones.

Type identity is the most obvious way to ensure that non-physical and physical causes do not comprise two distinct causes: if exactly the same particulars *and* properties comprise a non-physical and a physical cause, the two causes will certainly themselves be fully identical. Still, type identity is a very strong doctrine. Type identity about thoughts, for example, would imply that the property of *thinking about the square root of two* is identical with some physical property. And this seems highly implausible. Even if all human beings with this thought must be distinguished by some common physical property of their brains—which itself seems highly unlikely—there remains the argument that other life forms, or intelligent androids,

will also be able to think about the square root of two, even though their brains may share no significant physical properties with ours.

This 'variable realization' argument has led many philosophers to seek an alternative way of reconciling the efficacy of non-physical causes with the causal-closure thesis: one which does not require the strict identity of non-physical and physical properties. They suggest that it is enough if the non-physical properties *metaphysically supervene* on the physical ones. According to this doctrine, a being's physical properties will fix its non-physical properties, in the sense that any two beings who share physical properties will necessarily share the same non-physical properties, even though the physical properties which so 'realize' the non-physical ones can be different in different beings. This arguably ensures that nothing more is required for any specific instantiation of a non-physical property than its physical realization—even God could not have created my brain states without thereby creating my thoughts—yet avoids any reductive identification of non-physical properties with physical ones. (Those who hold that non-physical properties generally supervene on physical ones are therefore called 'non-reductive physicalists'. For discussion of this position, see the essays in Kim 1993.)

It is clear that the causal-closure thesis requires *at least* metaphysical supervenience. If non-physical properties do not so much as metaphysically supervene on physical properties, then non-physical causes will be manifestly independent of physical causes. (For example, my thoughts will be genuinely extra to my brain states, in the sense that my brain states could possibly have existed without my thoughts.) And, given the causal-closure thesis, this would mean that any physical effects of non-physical causes would have two quite independent causes, the non-supervenient non-physical cause, as well as the physical cause guaranteed by the causal-closure thesis. Such rampant overdetermination seems absurd. Sometimes specific events are overdetermined by two independent causes, each of which would have sufficed on its own—as when someone dies because they are shot and struck by lightning simultaneously. But it seems quite wrong to hold that *all* the physical effects of non-physical causes are similarly overdetermined by two independent causes.

Still, is metaphysical supervenience enough to avoid an implication of unacceptable overdetermination? Advocates of type identity insist that it is not. If non-physical properties merely supervene on physical properties, they are not the same properties, and the causes which involve these properties cannot be identical. If two beings can both be thinking about the square root of two, even though they share no physical properties, then their so thinking cannot be identical to their possession of any physical property. Advocates of type identity thus argue that supervenience without identity still leaves the physical effect of any non-physical cause with two distinct causes, the physical cause guaranteed by the causal-closure thesis, and the distinct non-physical cause, and this itself is unacceptable (see Kim 1993).

The issue here hinges on the acceptability of different kinds of overdetermination. All can agree that 'strong overdetermination' is absurd—the physical effects of non-physical causes surely don't always have two quite independent causes—and

thus that causal closure implies at least the metaphysical supervenience of the non-physical on the physical cause. But this leaves open the possibility of 'weak overdetermination', with the relevant physical effects always having both a physical cause and a supervenient but distinguishable non-physical cause. Opinions differ on the acceptability of weak overdetermination. Advocates of type identity argue that this is just as absurd as strong overdetermination. (What difference can the non-physical cause make, once the physical cause has done its work?) Friends of supervenience counter that there is nothing wrong with weak overdetermination. (In particular, it doesn't seem to carry the unhappy implication that the result would have occurred anyway, even if one of the causes had been absent—for the supervenience relation seems tight enough to ensure that both causes would have been absent if either were; see Bennett 2003.)

Given the plausibility of the arguments for variable realization, the majority of contemporary philosophers probably favour supervenience over type identity. Note how this considerably dilutes the ontological implications of the causal-closure thesis. It is one thing to hold that any causally efficacious non-physical property must be type identical with some physical property. It is far weaker to claim that any such non-physical property must supervene on physical properties. A wide range of different kinds of non-physical properties arguably satisfy the latter requirement. In particular, to return briefly to the methodological aspect of naturalism, it is by no means obvious that any property that supervenes on physical properties must be the kind of property that is open to scientific investigation. True, the most familiar kind of supervenient properties are *functional* properties, properties which are constituted by the role they play in some system of scientific laws, and these by their nature are appropriate objects of scientific investigation. But functional properties are by no means the only kind of properties that can be held to supervene on physical properties. For example, various kinds of properties that are constituted by normative facts are widely held to supervene on the physical, and it is doubtful that these are appropriate objects of scientific investigation.

Let me now turn to properties that are not themselves physically efficacious, and so escape the immediate impact of the causal-closure thesis. As observed in section 4.1, the causal closure of physics itself leaves it quite open that there should be non-physical realms which are causally isolated from the physical—a realm of ghostly spirits, perhaps. The argument from causal closure bites only on those items that do have physical effects, implying that they can only have those effects if they are themselves physical.

This now opens the possibility that there are features of reality that do not even supervene on the physical realm, notwithstanding the causal closure of physics. There are two philosophically significant options to consider under this heading: first, epiphenomenal mental kinds; second, commonsensically non-causal kinds like mathematics, modality, and morality.

The first option relates specifically to *conscious* mental states. Pretheoretically, these might be supposed to have physical effects (most obviously of a behavioural kind) and therefore to supervene on the physical realm. But many philosophers feel that

there are compelling reasons to deny that conscious properties are metaphysically supervenient on the physical. (Aren't zombies metaphysically possible—beings who share all your physical states yet have no conscious life at all?) Given causal closure, this non-supervenience threatens to imply strong overdetermination of all the physical effects of conscious causes—unless, that is, we conclude that conscious causes don't have physical effects after all. This epiphenomenalism option has been explored by various contemporary philosophers. Persuaded that conscious properties can't possibly be physical, they conclude that conscious states must be causal 'danglers', causal upshots of brain processes, but without any effects themselves (see Jackson 1982; Chalmers 1996). (Other philosophers, perhaps the majority, run the argument the other way, sticking to the claim that conscious properties have behavioural effects, and concluding that they must therefore supervene on the physical, despite initial philosophical impressions to the contrary.)

Now consider mathematical, modal, and moral facts. Here there is plenty of room from the start to doubt that they have physical effects. Mathematical entities are abstract, outside space and time; modal facts concern realms of possibility other than the actual; moral facts seem to transcend any natural facts. If it follows from such considerations that these facts have no physical effects, then they escape the argument for physicalism. Causal closure argues that anything with a physical effect must itself be physical. But this carries no immediate implications for mathematical, modal, or moral facts, if they lack physical effects. For all that has been argued so far, they need have no particular relation to the physical world at all.

The idea of facts without physical effects does face other difficulties, however. In particular, there are obvious epistemological queries about such facts. If mathematical, modal, and moral facts—or, for that matter, epiphenomenal conscious facts—lack any physical effects, then how can we possibly know about them? The standard route to knowledge of the actual world is via perception, or at least via some kind of causal interaction between our cognitive systems and the facts known about. And since our cognitive systems are physically realized, this would seem to require that the facts known about have physical effects. This makes it hard to see how we can know about realms of reality which lack physical effects.

This argument is not conclusive. There remains the option that we can know about these realms a priori, or in some other way that does not require causal interaction. A number of contemporary philosophers defend theories along these lines. Still, it is not always obvious that such non-causal epistemologies can be made to work. An alternative response to the epistemological difficulties is to conclude that the putative non-causal entities do not really exist, and that our discourse about them is no more than a useful fiction. In addition, there remains the option of insisting that the problematic area is metaphysically supervenient on the physical after all, and so capable of interacting physically with our cognitive systems.

Let me conclude by summing up this final section. The causal-closure thesis argues that anything with physical effects must in some sense be physical. There may be room in reality for non-physical things that lack physical effects, but it is not obvious that we can have knowledge of such things. However, even if our knowledge

is restricted to physical things, the requirement of physicality is not necessarily as demanding as it may seem. Some take it to require type identity between physical and prima facie non-physical properties. But as many hold that it only requires the metaphysical supervenience of the prima facie non-physical on the physical, and this is a far weaker requirement.

References

Bennett, K. (2003), 'Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It', *Noûs*, 37: 471–97.

Chalmers, D. (1996), *The Conscious Mind: in Search of a Fundamental Theory* (Oxford: Oxford University Press).

Crane, T., and Mellor, D. (1990), 'There Is No Question of Physicalism', *Mind*, 99: 185–206.

Hempel, C. (1980), 'Comments on Goodman's *Ways of Worldmaking*', *Synthese*, 45: 193–99.

Jackson, F. (1982), 'Epiphenomenal Qualia', *Philosophical Quarterly*, 32: 127–36.

____ (1998), *From Metaphysics to Ethics: A Defence of Conceptual Analysis* (Oxford: Oxford University Press).

Kim, J. (1993), *Supervenience and Mind* (Cambridge: Cambridge University Press).

Lowe, E. (2000), 'Causal Closure Principles and Emergentism', *Philosophy*, 75: 571–85.

McLaughlin, B. (1992), 'The Rise and Fall of British Emergentism', in A. Beckerman, H. Flohr, and J. Kim (eds.), *Emergence or Reduction* (Berlin/New York: de Gruyter, 49–93).

Papineau, D. (2002), *Thinking About Consciousness* (Oxford: Oxford University Press).

Pettit, P. (1993), 'A Definition of Physicalism', *Analysis*, 53: 213–23.

Woolhouse, R. (1985), 'Leibniz's Reaction to Cartesian Interaction', *Procedings of the Aristotelian Society*, 86: 69–82.