dialectica

dialectica Vol. 64, N° 3 (2010), pp. 363–384
DOI: 10.1111/j.1746-8361.2010.01237.x

Closure Principles and the Laws of Conservation of Energy and Momentum

Sophie Gibb[†]

Abstract

The conservation laws do not establish the central premise within the argument from causal overdetermination – the causal completeness of the physical domain. Contrary to David Papineau (2000 and 2002), this is true even if there is no non-physical energy. The combination of the conservation laws with the claim that there is no non-physical energy would establish the causal completeness principle only if, at the very least, two further causal claims were accepted. First, the claim that the only way that something non-physical could affect a physical system is by (1) affecting the amount of energy or momentum within it, or (2) redistributing the energy and momentum within it. Second, the claim that redistribution of energy and momentum cannot be brought about without supplying energy or momentum. Both of these claims, however, are exceedingly difficult to defend in the context of the argument.

Despite the importance of the principle of the causal completeness of the physical domain to physicalism, rigorous argument for it is hard to find – presumably because it is assumed to be an uncontentious claim that is supported by physics and that therefore requires little further defence from those in the mental causation debate. But how, exactly, is this principle supported by physics?

One might presume that an answer to this question is to be found in the laws of conservation of energy and momentum – a cornerstone of contemporary physics. The thought that interactive dualism clashes with the conservation laws dates back to early criticisms of Descartes's theory of psychophysical causation, and is still popular in today's mental causation debate.[1] As Dennett writes:

> the principle of the conservation of energy is apparently violated by dualism. This confrontation between quite standard physics and dualism has been endlessly discussed since Descartes' own day, and is widely regarded as the inescapable and fatal flaw of dualism (1991, 35).

It is therefore perhaps unsurprising that it is these laws from which some assume the completeness principle ultimately to gain its support.

[†] Department of Philosophy, Durham University, 50 Old Elvet, Durham, DH1 3HN, UK; Email: s.c.gibb@dur.ac.uk

[1] For examples of those who draw interactive dualism into question on the basis of the conservation laws, see Churchland (1994, 20), Dennett (1991, 35), Flanagan (1991, 21), Fodor (1994, 25), Harbecke (2008, 24), Heil (1998, 23), Hopkins (1978, 223), Levine (2001, 5), Papineau (2000 and 2002), Pollock (1990, 19) and Searle (2004, 42).

Published by Blackwell Publishing Ltd., 9600 Garsington Road, Oxford, OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

Unfortunately, those who appeal to the conservation laws to attack interactive dualism seldom offer anything in the way of detailed argument, assuming the inference from the conservation laws to the rejection of interactive dualism to be a straightforward one. David Papineau (2000 and 2002) – a leading proponent of the completeness principle – is a notable exception and one of the few to explore in detail how one might attempt to construct an argument for the completeness principle on the basis of the conservation laws. Papineau (2000, 185) recognises that such an argument will be far from straightforward, arguing that it requires the premise that there is no non-physical energy.[2] However, his central thesis is that over the last century new empirical evidence has arisen which provides an inductive argument against the existence of non-physical energy. In virtue of this, he thinks that one can conclude that the completeness principle "by any normal inductive standards, has been fully established" (Ib., 203).[3]

In this paper, I aim to establish that even in combination with the claim that there is no non-physical energy, the conservation laws do not establish the completeness principle.[4] I shall argue that, at the very least, two further causal claims are also required. First, the premise that the only way that something non-physical could affect a physical system is by (1) affecting the amount of energy or momentum within it, or (2) redistributing the energy and momentum within it. Second, the premise that redistribution of energy and momentum cannot be brought about without supplying energy or momentum. Unlike the premise that there is no non-physical energy, for which Papineau has arguably demonstrated that there is empirical evidence, I shall suggest that plausible non-question-begging arguments for these additional premises are hard to find.

[2] Also see the appendix of Papineau's *Thinking about Consciousness* (2002, 234). Note that, as Papineau comments in the preface of *Thinking about Consciousness*, most parts of the appendix descend from 'The Rise of Physicalism'. Consequently, the criticisms raised against Papineau (2000) within this paper are equally criticisms against Papineau (2002).

[3] Although Papineau's discussion is framed in terms of forces rather than in terms of energy, to be consistent with those such as Hart (1988) who talk about psychic energy, and to avoid unnecessary complications, I shall interpret Papineau's argument as an argument against the existence of sui generis non-physical energy. This is not to misrepresent Papineau's position. In classical contexts, force-based formulations of mechanics and energy-based formulations are arguably interderivable: the terms (and principles) of either theory can arguably be derived from those of the other. Hence, energy (both kinetic and potential) is initially defined in terms of the work done by a force acting on a body, so potential and kinetic energy can be derived from force, and equally, force can be derived from potential energy. Given that sui generis mental energy can be defined in terms of the work done by a sui generis mental force, evidence against mental forces can also be taken to be evidence against mental energy. (See, for example, Goldstein et al., 2002, section 1.1).

[4] For support of the claim that Papineau is wishing to combine the claim that there is no non-physical energy with the conservation laws to establish the completeness principle, see section 1.

If my claims are correct, they lead to the rejection of Papineau's argument for the completeness principle. More generally, they raise a serious problem for all of those within the mental causation debate who would wish to reject interactive dualism via an appeal to the conservation laws and serve to dispel any idea that the inference from the conservation laws to the rejection of interactive dualism is an unproblematic one.

1. The argument from causal overdetermination and the conservation laws

Let me begin by outlining the argument from causal overdetermination. Upon consideration of the relation that we have with the physical world, it seems wholly apparent that some mental causes have physical effects. The argument from causal overdetermination combines this premise of psychophysical causation with two further premises, to yield the conclusion that such mental causes are identical with physical causes:

1. Mental causes have physical effects.[5]
2. Every physical effect has a sufficient physical cause (*Completeness*).[6]
3. There is no systematic causal overdetermination.

Mental causes (that have physical effects) are identical with physical causes.

In brief, given *Completeness*, if a physical effect has a non-physical cause, it would be *systematically* (that is, as a general rule) causally overdetermined, contrary to the third premise. To avoid denying the third premise, while accepting psychophysical causation, one must therefore identify those mental causes that have physical effects with physical causes.

Clearly, the argument from causal overdetermination raises a problem for substance dualism – if mental substances are non-physical, then given *Completeness* and the denial of systematic causal overdetermination, the mental must be wholly epiphenomenal in the physical domain. But, as is now generally recognised, psychophysical causation is in fact problematical for *any* position that maintains that mental and physical properties are distinct, regardless of whether these properties belong to the same substance. Given any property dualism, the question

[5] I take the causal relata to be Kimean events, that is to be the instantiation of a property by a substance at a time. However, the central arguments of this paper would not be affected if one were to advance an alternative account of the causal relata.

[6] To allow for the possibility of indeterministic causation, *Completeness* should be modified to the claim that: 'Every physical effect has a set of physical causes that together are sufficient to fix the chances of its occurrence'. I shall, however, follow Papineau (as it is his position that is my central concern) in ignoring probabilistic versions of *Completeness*, which he claims "would only complicate the issues unnecessarily" (2000, 203, fn. 2).

of whether mental properties are causally efficacious in the physical domain arises. If one were to deny mental properties such causal efficacy, this would be to abandon any serious commitment to psychophysical causation. On the other hand, to allow them such causal efficacy, given the denial of systematic causal overdetermination, would arguably be to reject *Completeness*.

For this reason, the argument from causal overdetermination raises, rather ironically (for one of the main advantages of physicalism was thought to be its ability to accommodate psychophysical causation), a serious problem for physicalism. This is because non-reductive physicalism – a position which is, in Crane's (1995) words, the 'orthodox' form of physicalism – is a type of property dualism. The reason why this is the orthodox form of physicalism is that physicalists generally consider the alternative forms of physicalism to be unsatisfactory. Psychophysical reductionism, which identifies mental properties with physical properties, is thought to fail because of the argument from multiple realisability. According to this argument, mental properties are multiply realised by, and hence cannot be identical with, physical properties. Alternatively eliminativism, which rejects mental properties and hence mental causes, is not widely accepted by physicalists because most would not wish to deny the common-sense claim that there is psychophysical causation.

Whether the non-reductive physicalist can respond to the argument from causal overdetermination by appealing to the dependence relationship that he understands to exist between mental and physical properties is debatable.[7] The response that is not available to the non-reductive physicalist, or indeed any physicalist – for this would be to abandon physicalism itself – is to reject *Completeness*.[8] However, physicalism is not sacrosanct, and given the undesirable state of affairs that the argument from causal overdetermination leaves us with, it is surely this central premise of the argument to which attention should turn.

Now the popular assumption within the mental causation debate is that *Completeness* is a fact of current physics. However, if *Completeness* is a working hypothesis of current physics, then it is one that is left wholly implicit – the principle is not referred to in any physics textbook.[9] Some argument is therefore needed for the claim that *Completeness* is a fact of current physics. In his discussion of the completeness principle, Papineau (2000 and 2002) attempts to provide

[7] Yablo (1992, section 4), Papineau (1993, 25), Jackson and Pettit (1988, 391) and Dretske (1988, ch. 2) advance different versions of this defence. For criticisms of these kinds of response, see Kim (1992, 1993a, 1993b).

[8] For defence of the claim that to reject *Completeness* is to reject physicalism, see Crane (2001, 43–48).

[9] For further support of the claim that *Completeness* is not referred to in any physics textbooks, see Papineau (2000, 184–185).

such an argument. His initial idea was that *Completeness* is a straightforward consequence of the laws of conservation of energy and momentum (Papineau 2000, 185). According to these laws:

> *Conservation*: Every physical system is conservative or is part of a larger system that is conservative (where a system is conservative if its total amount of energy and linear momentum can be redistributed, but not altered in amount, by changes that happen within it).[10]

Given *Conservation*, one might argue, *Completeness* must be true – if non-physical mental events did affect neurophysiological events, this would alter the total energy and/or linear-momentum of the brain, thereby violating *Conservation*. However, the striking problem with this argument is, Papineau observes, that the conservation laws do not tell us which forms of energy there are, only that they must be conserved (Ib., 196). Consequently, if there were such a thing as sui generis mental energy, then, as long as the total energy is conserved, mental energy could bring about physical effects without violating *Conservation*. But Papineau addresses this problem, providing a detailed argument to suggest that, in the light of recent evidence from theoretical physics and physiological research, there is probably no non-physical energy.[11] According to Papineau, this establishes that *Completeness* is probably true.

Now the issue of whether there is non-physical energy is a crucial one to the mental causation debate. Dualists such as Hart (1988) have appealed to 'psychic' energy in order to advance a theory of psychophysical interactionism. If Papineau's claim that it is implausible that there is non-physical energy is correct, then these theories of psychophysical causation must be abandoned. However, the question that I wish to consider is not whether Papineau is correct to maintain that current physics allows one to advance an argument against the existence

[10] For this formulation of the conservation laws, see Broad (1925, 105). I use this formulation as it is the kind that Papineau is arguably assuming in his discussion. Note that this formulation coincides with the one given in the *Oxford Dictionary of Physics* (Daintith 2005), which states that a conservation law is a law stating that the total magnitude of a certain property of a system, "such as its mass, energy, or charge, remains unchanged even though there may be exchanges of that property between components of the system". I limit my discussion to the laws of conservation of energy and momentum, as Papineau's discussion is itself limited to these conservation laws. Both Papineau's arguments and my own could be re-phrased to take into account other conserved quantities, but to do so would not significantly alter the dispute.

[11] More precisely, Papineau considers current science to provide two arguments against the existence of such forms of energy: (1) "The argument from fundamental forces" which claims that, on the basis of empirical evidence, we can inductively reason that "all apparently special forces characteristically *reduce* to a small stock of basic physical forces which conserve energy" (2000, 197–198 and 2002, 250); and (2) "The argument from physiology" which consists in the consideration that "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" (2000, 202 and 2002, 254).

of non-physical energy, but rather whether this does, as he claims, establish the probable truth of *Completeness*.

Let us call the premise that there is (probably) no non-physical energy '*Energy*'. Obviously, by itself, *Energy* does not entail *Completeness*, and hence must be supplemented with further premises. Nor, of course, is Papineau suggesting otherwise. It is quite clear from his discussion that he considers it to be *Energy* in combination with *Conservation* that provides the basis for an argument for *Completeness*. (After all, what prompted his discussion of whether there is non-physical energy was precisely the thought that if there isn't, then *Conservation* leads to *Completeness*.)

However, as it stands the following argument is clearly not satisfactory:

1. Every physical system is conservative or is part of a larger system that is conservative (*Conservation*).
2. There is (probably) no non-physical energy (*Energy*).

Every physical effect has a sufficient physical cause (*Completeness*).

First, a revision is required concerning the kind of causal principle that the combination of *Conservation* and *Energy* is being taken to provide evidence for. This revision is not the primary concern of this paper, and I shall therefore provide only a very brief discussion of it here.[12]

According to *Completeness*, the physical domain is causally complete. One can, in other words, tell the whole causal story about any effect in the physical domain, without ever having to leave this domain. Mental causes are therefore never *needed* to account for physical effects. This is not to rule out the possibility that mental causes *could* have physical effects. However, in every such case, the relevant effect would also have a complete physical cause, and hence the mental cause would be redundant.

There is, however, a second kind of causal argument that appeals to a causal principle of a much stronger type:

1. No physical effect has a non-physical cause.[13]
2. Mental causes have physical effects.

Mental causes that have physical effects are identical with physical causes.

This argument, unlike the argument from causal overdetermination, does not need the premise of the denial of systematic causal overdetermination. This is because

[12] For a detailed discussion of this issue, see Gibb (forthcoming).
[13] This principle is, for example, assumed by Smith and Jones (1986, 66).

its causal principle (premise 1), unlike *Completeness*, *directly* rejects the possibility of non-physical causes having physical effects. As this principle excludes the non-physical from having physical effects, I shall refer to it as an exclusion principle, or *Exclusion* for short.[14]

Now, *insofar* as the combination of *Conservation* and *Energy* supports either causal principle, it supports *Exclusion*.[15] Papineau's claim is that dualist theories of psychophysical causation violate the conservation laws, unless there is non-physical energy. Hence, the threat to dualism is not merely that the combination of *Conservation* and *Energy* leaves mental causes redundant in the physical domain, but that it excludes them altogether.

Note, however, if *Conservation* and *Energy* do provide support for *Exclusion*, strictly speaking it is arguably not incorrect to appeal to these premises to support *Completeness*. This is because *Exclusion* arguably entails *Completeness*. If so, these premises support *Completeness* in virtue of the fact that they support *Exclusion*.[16] However, the only reason for inserting the premise of the denial of systematic causal overdetermination within a causal argument is that it supposedly allows one to advance a weaker causal principle, and hence one that has a greater chance of empirical support. If, as Papineau's argument indicates, the best evidence for *Completeness* is in fact evidence for *Exclusion*, then there is no motivation for supporting the argument from causal overdetermination, as opposed to a strong causal argument containing only two premises. The contemporary mental causation debate would have to alter in order to reflect this, for it would entail that those discussions in the debate over the premise of the denial of systematic causal overdetermination would be redundant.[17]

2. *The hidden premises in the argument from Conservation*

Having noted this revision to the conclusion of Papineau's argument, we are now in a position to consider whether the combination of *Conservation* and *Energy*

[14] This is not to be confused with Kim's (2005) exclusion principle which refers to the principle of the denial of systematic causal overdetermination.

[15] See Gibb (forthcoming) for a detailed defence of this claim.

[16] Note, however, that *Exclusion* entails *Completeness* only if causation is deterministic. (That causation is deterministic is, as noted, an assumption of this paper.) Moreover, an anonymous member of the editorial committee has commented that if a Laplacian as opposed to a causal determinism is accepted, *Exclusion* still does not entail *Completeness*. This is a conclusion that I'm happy to accept for the purpose of this paper. The criticism of Papineau's argument that I go on to advance in no way depends on the claim that *Exclusion* entails *Completeness*.

[17] In Gibb (forthcoming), I argue that the move to a two-premise causal argument raises a serious problem for non-reductive physicalism.

really does provide an argument for *Exclusion*. Unfortunately, the following argument is still not valid:

1. Every physical system is conservative or is part of a larger system that is conservative (*Conservation).*
2. There is (probably) no non-physical energy (*Energy*).

(Probably) no physical effect has a non-physical cause (*Exclusion*).

This is because although the conclusion is a claim about causation, the premises make no reference to causation. Hence, in order to move from the combination of *Conservation* and *Energy* to *Exclusion*, Papineau must be implicitly assuming some claim about causation that enables one to relate energy to causation. Indeed, I would suggest that *two* additional substantive causal claims are required, without which neither *Conservation* nor *Energy* nor their combination could be used to defend *Exclusion*. I shall discuss each of these causal premises in turn.

The first causal premise that is required limits the ways in which a physical system can be affected. *Conservation* states that every physical system is conservative or is part of a larger system that is conservative, where a system is conservative if its total amount of energy and linear momentum can be redistributed, but not altered in amount, by changes that happen in it. *Conservation* is therefore inconsistent with a cause affecting a physical system by altering the amount of energy or momentum in it, unless that cause is from within a larger conservative system of which the first is a part. On the other hand, *Conservation* is consistent with a cause affecting a physical system by redistributing the amount of energy or momentum in it. Consequently, the following causal premise is needed in Papineau's argument:

> *Physical Affectability*: The only way that something non-physical could affect a physical system is by (1) affecting the amount of energy or momentum within it, or (2) redistributing the energy and momentum within it.

Conservation disallows (1) unless a physical system is part of a larger system that is itself conservative, but allows (2). If, contrary to *Physical Affectability*, there was a third way that something non-physical could affect a physical system, *Conservation* would have nothing to say about it, and therefore could not be appealed to in order to rule out the non-physical affecting a physical system in such a way.

Physical Affectability would appear to be an implicit assumption in Papineau's discussion, for he only considers the possibility that mental causes might affect the physical domain by "altering the physical positions of particles" or, in other words, by "moving matter" (2000, 189). To suggest that psychophysical causation must be like this is to assume *Physical Affectability*, for if all such physical changes are

ultimately changes in the motion of matter, then all such changes must be explicable in terms of energy and momentum redistribution. Note that Papineau is certainly not alone in accepting *Physical Affectability*. An inspection of the mental causation debate reveals that, in many cases, the Cartesian assumption that for there to be psychophysical causation mental causes must in some way bring about changes in the motion of matter, has survived the dualism of which it was once a part.[18]

Of course, one might reasonably wonder what a denial of *Physical Affectability* could consist in. However, one only has to turn to the current mental causation debate to supply an answer to this question. Some dualists in this debate consider that their theory of psychophysical causation depends upon the rejection of precisely this kind of premise. In particular, I have in mind E. J. Lowe (1993, 1996, 1999, 2000), a central proponent of substance dualism.

Lowe suggests that mental events must have independent causal powers because it is only if some physical effects have mental causes that we can:

> represent the physical effect of certain physical causes – such as a bodily movement produced by a large number of apparently independent neural events in the brain – as being non-coincidental, since there might be no identifiable *physical* event(s) which could be seen to link the chains of physical causation involved in an appropriate way (2000, 584).

Let me explain this position in more detail. As we trace the causal chains of neural events that results in a bodily movement backwards into the brain, Lowe observes that it seems likely that these causal chains will display a highly complex, fractal, tree-like structure. He argues that from a purely physical perspective, the fact that the causal tree of physical events converges upon a particular event – the bodily movement – looks purely coincidental, for as physical science "traces back the physical causes of our bodily movements into the maze of antecedent neural events, it seems to lose sight of any unifying factor explaining why those apparently independent causal chains of neural events should have converged upon the bodily movements in question" (Ib., 581). From this purely physical perspective, this convergence seems no less of a remarkable coincidence than if the rings on the surface of a pond were to converge upon a central point (Lowe 1996, 68). Lowe maintains that this convergence of neural events requires explanation and that it is only by appealing to the mental that we can do this. It is the specific causal role of mental events to render the fact that a causal tree of neural events converge upon a particular bodily movement non-coincidental. Given the intentional nature of mental events – the fact that a mental event is directed upon the occurrence of a particular bodily behaviour – he considers mental events to be ideally suited to play this role. Consequently, in response to the physicalist objection that even if

[18] To give two examples, see Searle (1984, 92) and Taylor (1992, 22).

some event were required to play this causal role within the physical domain, one would never need to appeal to mental events to do so, Lowe is able to respond that it is precisely the special intentional nature of mental events that enables them to play this causal role.[19]

I do not wish to defend this theory here, but simply to observe that Lowe's account of psychophysical causation is a denial of *Physical Affectability*. To claim that the causal role of mental events within the physical domain is that of making non-coincidental certain physical events is neither to claim that mental events affect the amount of energy or momentum within a physical system nor the distribution.[20] Lowe's whole point is that if we concentrate our attention solely on the chains of energy or momentum transfer that take place between the various neural events, then we lose sight of the fact that their convergence is non-coincidental. Nor is it Lowe's claim that mental events play the role of making the convergence non-coincidental by transferring energy or momentum to a neural event or any set of neural events within the causal chain of neural events, for, according to Lowe's position, there is no individuable neural event or set of neural events in the maze of neural events that the mental event can plausibly be said to transfer energy or momentum to (Lowe 1999, 66).

Here let me clarify one point. For Lowe, mental events do not play their causal role by bringing about any single neural event or set of neural events in the causal tree, which would be the most obvious way for psychophysical causation to involve energy and momentum redistribution. Rather, according to his position, every event in the maze of neural events has an immediate *physical* cause that is wholly responsible for bringing it about. But still, one might point out that, according to Lowe, a mental event is causally responsible for the fact that the maze of neural events has this convergence characteristic (Lowe 1996, 67). The convergence of chains of neural events is, one might argue, itself a certain distribution of energy and momentum. Hence, for a mental event to be responsible for this convergence is for it to be responsible for a certain distribution of energy and momentum. Therefore, Lowe's account of psychophysical causation *must* in fact be a case of the mental affecting the physical by redistributing energy and momentum, and is therefore not a denial of *Physical Affectability*.

[19] As Lowe emphasises his proposal "is not compatible with a supervenience theory of mental states, for its implication is that mental states such as volitions or desires have genuinely independent causal powers – that they make a difference to how the world goes over and above any difference made by physical events" (1996, 68).

[20] That *Physical Affectability* is the particular premise that Lowe is wishing to reject in advancing his account of psychophysical causation has been confirmed by personal conversation with Lowe. See further his (1996, 63–64).

A response to this objection requires one to consider the notion of 'convergence' that Lowe is assuming. When sand is poured through a funnel, the funnel makes the sand converge into a narrow stream. When coins are placed in a coin-sorting machine, the machine makes coins that are of the same type converge in a single tube. One might interpret Lowe's claim that a mental event is responsible for the fact that a maze of neural events converge upon a specific bodily movement in an analogous way. Hence, one might interpret Lowe as claiming that a mental event is responsible for directing various electrochemical signals along certain neural pathways so that they converge upon a particular bodily behaviour, in the same way that the sides of the funnel direct the flow of the grains of sand so that they converge into a narrow stream. Indeed, the causal tree that Lowe describes might be thought to represent the movement of initially widely dispersed electrochemical signals along neural pathways, via synaptic connections, to a single location. If this was Lowe's claim, then it would not be a denial of *Physical Affectability* because it would entail that the mental causes the physical by redistributing energy and momentum – the sides of the funnel causing the grains of sand to alter their direction involves energy and momentum redistribution, and similar considerations would seem to apply if mental events somehow altered the flow of neural activity. However, this is not Lowe's claim for he is not using the term 'convergence' in this sense. Instead 'convergence' is to be understood as a formal property exhibited by some causal trees of events, which consists in the fact that a set of mutually independent causes together bring about a certain effect. In claiming that a mental event is responsible for the fact that a maze of neural events converge upon a particular bodily behaviour, his claim is that the mental event is responsible for this feature of the causal tree of neural events. That is, the mental event is responsible for the fact that a set of mutually independent neural events together bring about a certain bodily movement.[21] Given this understanding of convergence as a formal property of certain causal trees, Lowe would argue that it simply does not follow that the convergence of chains of neural events should be interpreted as a certain distribution of energy and momentum.[22] Whether this provides a wholly persuasive account of psychophysical causation is a further

[21] Note, therefore, the causal tree that Lowe describes simply represents the causal relations between various neural events and is not supposed to in some way represent the movement of electrochemical signals along neural pathways via synaptic connections. To talk about causal chains of neural events and to talk about the flow of electrochemical signals along neural pathways, via synaptic connections, is to talk about two different things. A causal chain of events is an abstract structure whereas the latter is a physical structure. Unless, perhaps, one maintains an energy transference theory of causation, there is no reason to think that the first maps onto the second, and hence that, for example, neural event chains follow synaptic connections.

[22] These clarificatory points about the notion of 'convergence' that Lowe is appealing to arise from personal correspondence with Lowe, to whom I owe my thanks.

question – my aim in discussing Lowe's position has simply been to draw attention to the fact that there are philosophically sophisticated accounts of mental causation that deny *Physical Affectability* which deserve to be taken seriously.

Even if *Physical Affectability* is accepted, the combination of this premise with *Conservation* and *Energy* still does not entail *Exclusion*. If, in accordance with *Energy*, a physical system is not part of a larger conservative system that includes non-physical energy, then the combination of *Conservation* and *Physical Affectability* entails that the only way that something non-physical could affect a physical system is by redistributing the energy and momentum within it. But the combination of *Conservation*, *Physical Affectability* and *Energy* still leaves a gap for psychophysical causation – mental events might redistribute energy and momentum without themselves supplying energy or momentum. Hence, consider Broad's response to those who believe that *Conservation* threatens dualism. He suggests that mental events could:

> determine that at a given moment so much energy shall change from the chemical form to the form of bodily movement; and they determine this, so far as we can see, without altering the total amount of energy in the physical world (1925, 109).

Broad's claim seems to be that mental events prompt the transfer of energy between physical events. However, they do not do so by transferring energy. For this reason, such a claim is consistent with *Conservation*. To rule out this form of psychophysical interactionism, a second causal premise is required. According to it:

> *Redistribution*: Redistribution of energy and momentum cannot be brought about without supplying energy or momentum.

Bringing all of these premises together, we have the following argument for *Exclusion*:

1. Every physical system is conservative or is part of a larger system that is conservative (*Conservation*).
2. There is no non-physical energy (*Energy*).
3. The only way that something non-physical could affect a physical system is by (1) affecting the amount of energy or momentum within it, or (2) redistributing the energy and momentum within it (*Physical Affectability*).
4. Redistribution of energy and momentum cannot be brought about without supplying energy or momentum (*Redistribution*).

No physical effect has a non-physical cause (*Exclusion*).

Given the complexity of this argument, any thought that one can move easily from the combination of *Conservation* and *Energy* to an anti-dualist conclusion is clearly incorrect. In addition to *Conservation* and *Energy*, *at least* two further

substantive claims are required, without which neither *Conservation* nor *Energy* nor their combination could be used to defend *Exclusion*. Furthermore, as I shall go on to argue in section 3, while Papineau has arguably provided a plausible argument for *Energy*, whether a plausible argument for either *Physical Affectability* or *Redistribution* is available is debatable.

Note that my concern is only to highlight two of the premises that the combination of *Conservation* and *Energy* must be supplemented with in order to reach *Exclusion*. I am not committed to the claim that these are the only additional premises that are needed and hence am not committed to the claim that the argument, even with the addition of these causal premises, is a valid one. I simply wish to draw attention to the unstated causal premises that such an argument must contain, premises that I consider to lack support, and hence that I consider, by themselves, to lead one to question whether one can build a persuasive argument for *Exclusion* via an appeal to the conservation laws.

To summarise: To move from the combination of *Conservation* and *Energy* to *Exclusion* additional causal premises are required. The further step of identifying what these causal premises are is done by considering what must be true, from a causal point of view, if *Conservation*, in combination with *Energy*, is taken to exclude physical effects from having non-physical causes. Since, given *Energy*, a physical system is not part of a larger conservative system that includes non-physical energy, if the non-physical affected a physical system by creating energy (or momentum) within it, this would violate *Conservation*, as the resulting energy gain would not be compensated by an energy loss. However, the combination of *Conservation* and *Energy* leaves open the possibility that the non-physical might affect a physical system by redistributing the energy and momentum within it, as long as this is not done by contributing energy or momentum to it. *Redistribution* is required to rule this possibility out. The combination of *Conservation*, *Energy* and *Redistribution* entails that the non-physical could not affect a physical system by affecting the amount of energy or momentum within it or by redistributing the energy or momentum within it. But what if there was some alternative way that the non-physical could affect a physical system? *Physical Affectability* is required to rule this out.[23]

[23] In 'Conservation of Energy is Relevant to Physicalism', Koksvik (2007) combines the conservation laws with premises that are different to those that I have outlined, in order to move from these laws to physicalism. One might therefore plausibly question whether I have provided a fair reconstruction of this argument, for Koksvik's argument unlike that of my own does not include *Physical Affectability* and *Redistribution*. Does this not show, therefore, that the physicalist can provide an argument that appeals to the conservation laws, whilst avoiding the acceptance of the troublesome causal premises? However, a closer examination of 3* of Koksvik's own argument establishes that this is not the case. According to 3*: "If a non-physical mind changes a physical system, it changes its energy level" (Ib., 579). 3* would only be true if: (1) The only way that a physical system could be causally affected by a non-physical mind is by

3. Defending Physical Affectability and Redistribution

To move from *Conservation* to *Exclusion* it is not enough, pace Papineau, to combine *Conservation* with *Energy*. Two causal premises, *Physical Affectability* and *Redistribution*, are also required. The plausibility of Papineau's argument for *Exclusion* therefore depends, in part, upon their plausibility. However, those that are sympathetic to Papineau's position might be unmoved by this, for they might consider that these additional hidden premises are wholly unproblematic and, indeed, take their truth for granted. In this final section, I would like to present some reasons for questioning this thought – upon closer inspection, the plausibility of *Physical Affectability* and *Redistribution* is dubious.

How might one attempt to defend *Physical Affectability* and *Redistribution*? It should immediately be made clear that the answer is not to be found in the energy transference theory of causation.[24] *Physical Affectability* and *Redistribution* are both entailed by this theory of causation, which claims that:

> *Transference*: Causation is the transference of a quantity from cause to effect, where this quantity is energy or momentum.

Transference entails *Physical Affectability*, for according to *Transference* all effects are energy or momentum gains. *Transference* entails *Redistribution*, for according to it all causation is energy or momentum transfer.

But in the context of the above argument, *Transference* cannot be appealed to to defend these premises. This is for the simple reason that if one replaces *Physical Affectability* and *Redistribution* with *Transference* in the above argument, it renders *Conservation* redundant. Such a replacement would leave us with the following simpler argument for *Exclusion*:

1. Every physical system is conservative or is part of a larger system that is conservative (*Conservation*).

affecting the amount of energy within it or redistributing it; and (2) Redistribution of energy cannot be brought about without supplying energy. Hence in order to accept 3* one must first accept *Physical Affectability* and *Redistribution*. Indeed, as Koksvik (Ib., 581) himself comments in the concluding section of his paper, an obvious way to reject his reconstructed argument is to reject 3*, so we would seem to be largely in agreement that it is this area of the physicalist's argument that needs closer examination. Indeed, it is precisely the goal of my paper to provide a clear formulation of the assumptions that lie behind the acceptance of something such as 3* and then to assess the plausibility of these assumptions. It should also be noted that the argument that Koksvik constructs makes no appeal to *Energy*, for it is instead articulated in terms of what Koksvik refers to as "e-closed systems" (Ib., 579). However, the specific aim of my paper has been that of establishing how, if Papineau is correct to maintain *Energy*, this might be appealed to in order to defend physicalism. It is therefore only correct that *Energy* appears as a central premise within my reconstructed argument. I am grateful to an anonymous referee for bringing Koksvik's paper to my attention.

[24] Proponents of this theory of causation include Aronson (1971), Fair (1979) and Hart (1988).

2. There is no non-physical energy (*Energy*).
3. Causation is the transference of a quantity from cause to effect, where this quantity is energy or momentum (*Transference*).

No physical effect has a non-physical cause (*Exclusion*).

This argument is unsatisfactory because *Conservation* is redundant within it – the combination of *Transference* and *Energy* on its own entails *Exclusion*. Clearly, if all causation is energy (or momentum) transfer, and all energy is physical, then all causes are physical.[25]

As an aside, one might think that the conclusion to be drawn from this is that Papineau's claim that there is no non-physical energy should be combined with *Transference* rather than *Conservation* to get to *Exclusion*. That is, one should turn from an appeal to the conservation laws (which combines *Energy* with *Conservation*, *Physical Affectability* and *Redistribution* to establish *Exclusion*), to an appeal to the energy transference theory of causation (which combines *Energy* with *Transference* to establish *Exclusion*) – the resulting redundancy of *Conservation* is of no consequence to his physicalist stance. But this will not work, for what is of concern is whether independent support can be provided for *Transference* and there is reason to doubt that it can be. The energy transference theory of causation takes physical relations as its paradigm for understanding causal relations – the examples that Fair (1979), Aronson (1971) and even Hart (1988) (a substance dualist) give to motivate this theory of causation are all based in the physical domain. Consequently, it is arguable that in so far as it provides a satisfactory account of causation, it is one that applies to the physical domain only. Now, unless one denies the homogeneity of the causal relation, a satisfactory theory of causation must provide a plausible account of *all* causation, not just causation in a particular domain. Therefore, if, as is most plausible, the energy transference theory is a theory about *physical* causation only, for it to provide a satisfactory theory of causation, independent support needs to be found for the claim that all causation is physical causation. But supposedly the best argument for this claim is the causal argument. Hence, to support *Transference*, one must first assume the

[25] Regarding this point, I am therefore in agreement with Montero (2006). Note, however, that Montero's claim that the conservation laws have nothing to do with physicalism is too hasty. This claim rests upon her initial assumption that, to be valid, an argument for physicalism via the conservation laws *must* incorporate the energy transference theory of causation as a premise. This assumption is false. Montero is assuming a causal claim that is stronger than is necessary to move from the conservation laws to *Exclusion*. That is, to move from the conservation laws to *Exclusion*, one does not require *Transference* – the weaker combination of *Physical Affectability* and *Redistribution* is sufficient. And, with this weaker combination, *Conservation* is not rendered redundant.

causal argument and hence the truth of *Exclusion*. However, to support *Exclusion*, via the appeal to the energy transference theory of causation, one is assuming *Transference*.[26]

Let us therefore return to our original argument which combines *Energy* with *Conservation*, *Physical Affectability* and *Redistribution* in an attempt to establish *Exclusion*. Our question was what reason is there to support the premises of *Physical Affectability* and *Redistribution* in this argument? We have seen that *Transference* cannot be appealed to in order to supply an answer. But is there any reason to support *Physical Affectability* and *Redistribution* independently of *Transference*? Rather than appealing to some metaphysical fact about causation, one might presume that, as with *Energy*, *Physical Affectability* and *Redistribution* can be established by some empirical fact about the physical domain – philosophers such as Lowe and Broad, who consider that the most plausible solution to the problem of mental causation is to reject either *Physical Affectability* or *Redistribution*, are failing to take this empirical fact into account.

It is, however, far from clear what this empirical fact might be. As with *Completeness*, if *Physical Affectability* and *Redistribution* are working hypotheses of physics, then they are left wholly implicit. Neither principle is written down in any physics textbook. Nor – a point to which I shall return – does the rejection of either principle lead to the violation of any physical law. It should also be noted that we cannot turn to Papineau's discussion for an answer. Although Papineau has provided a detailed empirical defence for *Energy*, no argument for either *Physical Affectability* or *Redistribution* is to be detected in his writing. It is perhaps worth pointing out that Papineau's Argument from Fundamental Forces and his Argument from Physiology – the two arguments that he provides for *Energy* – do not provide any support for *Physical Affectability* or *Redistribution*. To remind the reader, the Argument from Fundamental Forces is the argument that we have good empirical evidence that all apparently special forces reduce to physical forces and the Argument from Physiology is the argument that if special forces existed, physiological investigation would reveal their presence, but it does not.[27] First, consider *Physical Affectability*. According to Newtonian Law, the effect of a force is to bring

[26] Indeed, note that if the energy transference theory of causation is to be interpreted as presenting a plausible theory of causation, then unlike *Conservation*, which can be maintained independently of *Energy*, it is arguable that *Transference* cannot be. Fair (1979, 222–223) argues that in order to provide an informative definition of the term 'energy', and hence for the energy transference theory of causation to provide an interesting theory of causation, energy must be physical. This is because it is only within physics that the term 'energy' has a sense in which it is not synonymous with the term 'cause'. If Fair is correct, the only plausible version of the energy transference theory of causation assumes *Energy*, and hence entails *Exclusion*. Consequently, to reconcile *Transference* with the claim that there is mental causation, psychophysical reductionism must be correct.

[27] See further fn. 11.

about proportional changes in the velocities of the bodies it acts on. To deny *Physical Affectability*, that is, to deny that the non-physical affects a physical system by affecting the amount or distribution of energy or momentum within it, is to deny that all physical change is change in the velocities of bodies. Hence, it is to deny that all causes are forces or, more precisely, that all causes involve the exertion of a force – a dualist who denies *Physical Affectability* will not understand psychophysical causation in terms of pushes or pulls. As a dualist who rejects *Physical Affectability* is not committed to the claim that mental causes are forces, considerations about whether mental forces reduce to physical forces or whether there is any physiological evidence for mental forces are beside the point. Similar considerations apply with regard to *Redistribution*. By definition, whenever a force is exerted upon an object, this involves the supply of energy or momentum. (It must not be forgotten here that, from a theoretical perspective, force-based and energy-based formulations of mechanics are mutually supporting, interderivable accounts of mechanical motion.) To deny *Redistribution* is to maintain that an entity can cause the redistribution of energy and momentum without supplying energy or momentum. Hence, it is to maintain that an entity can cause this redistribution without exerting a force. Thus to deny *Redistribution* is not to assume anything about special forces and therefore, as with *Physical Affectability*, Papineau's two arguments are not relevant in assessing its plausibility.

Admittedly, Papineau does briefly criticise Broad's theory (which denies *Redistribution*) in a footnote, claiming that "it is difficult to avoid the impression that he has mastered the letter of the principle of the conservation of energy, without grasping the wider physical theory in which it is embedded" (2000, 206, fn. 20). This point is not developed. Elsewhere, it has been argued that Broad's theory is inconsistent with the law of conservation of momentum and it might be this that Papineau is referring to (Cornman 1978). This point may, in turn, lead one to question whether any denial of *Redistribution* could be consistent with this law. However, as Averill and Keating (1981) forcefully argue, this criticism of Broad rests upon an implausibly strong formulation of the conservation laws. Furthermore, the denial of *Physical Affectability* is certainly not inconsistent with this law or indeed any other law of physics.

To provide empirical support for the latter premise, one might instead resort to the following kind of inductive argument. At the fundamental level all change can be accounted for in terms of energy and momentum redistribution. No other kind of change has been observed amongst physical particles, nor does physics provide us with any reason to think that any other kind of physical change might exist. This gives us good reason to think that *all* physical change is fundamentally nothing but changes in energy or momentum distribution. Hence, given the evidence from within current physics, *Physical Affectability* is most probably correct. I do not think, however, that this argument would worry those interactive dualists who

would wish to deny *Physical Affectability.* Their whole point is that bodily behaviour that is caused by propositional attitudinal states is, in certain respects, importantly different from other physical effects – that when it comes to these physical effects, to provide a full causal account of them we must appeal to something other than energy and momentum redistribution. Hence, consider Lowe's theory, which is a denial of *Physical Affectability*. According to it, what sets such bodily behaviour apart from other physical effects is that it is caused by a large number of independent physical events where this convergence is most plausibly non-coincidental. And, as discussed in the previous section, Lowe considers that the fact that this convergence is non-coincidental cannot be explained in terms of energy and momentum redistribution.

Might it instead be argued that *Physical Affectability* should be accepted because the deductively closed set of propositions that provides the best balance of explanatory strength and simplicity, and that incorporates the best contemporary physics is the set that includes it?[28] First, let it be noted that to reject *Physical Affectability* is not to reject the best contemporary physics. As commented, *Physical Affectability* is neither a working hypothesis of contemporary physics nor a consequence of any of its laws. Furthermore, it is debatable whether the system that provides the best balance of explanatory strength and simplicity is one that accepts *Physical Affectability*. Once again consider Lowe's position. According to him, a system that denies *Physical Affectability* will have greater explanatory strength than a system that accepts it, for only the former system will be able to explain the convergence of neural events.

There is one final method by which one might attempt to defend *Physical Affectability* (and, perhaps, *Redistribution*). This is to argue that any denial of *Physical Affectability* is implausible because it will render psychophysical causation unintelligible. Such an objection might be expressed in the following way: those dualists who deny *Physical Affectability* can provide no plausible response to the question of *how* mental events cause physical events. Thus, regarding Lowe's theory, how do mental events bring about the fact that there exists a neural maze with a particular convergence characteristic, if not by redistributing energy and momentum? Lowe's account lacks intelligibility, because no alternative account is offered of *how* mental events actually do this.

To respond to this objection, in asking for an explanation of *how* mental events render physical events non-coincidental, or more generally, in asking for an explanation of *how* a mental event causes a physical event, one is inquiring about the causal mechanism behind psychophysical causation. To demand a causal mechanism, that is, to consider the question of how a cause brings about its *direct* effect to be a legitimate one, is to make certain assumptions about the nature of the causal

[28] I am grateful to a member of the Editorial Committee for this suggestion.

relation. The regularity theory of causation is, for example, non-mechanistic. It maintains that C causes E if and only if C and E are constantly conjoined. There is no further fact about the cause or effect, or about the relevant sequence that gives it its causal character. Thus, where C is a direct cause of E, the question of what it is about this constantly conjoined sequence that explains *how* a cause is able to bring about its direct effect simply has no application. The counterfactual theory of causation, those nomological theories of causation that provide a non-reductive analysis of a law, and Harré and Madden's (1975) account of causation which argues for the locus of causal powers in the 'powerful particular' are also non-mechanistic in this sense.[29] Given these theories of causation, the question of *how* mental events render physical events non-coincidental is simply not allowed. Thus, for example, taking Harré and Madden's account of causation and applying it to Lowe's theory of psychophysical causation, mental events have the causal power to make a causal tree of neural events converge upon a particular bodily movement, and they have this causal power in virtue of their intentional nature. To then demand an account of *how* mental events do this is to abandon this account of causation for a mechanistic one.

The demand for a causal mechanism, and hence the demand for an answer to the question of how, for example, mental events render physical events non-coincidental, is acceptable only if one advances a theory of causation that analyses causation in terms of underlying non-causal processes associated with causation – processes that can then be appealed to in order to explain how a cause brings about its *direct* effect. The energy transference theory of causation provides the best example of this type of theory of causation. According to it, the mechanism of causation is energy transfer. In answer to the question of how a cause brings about its direct effect one can respond that it is by transferring energy to the effect. However, as we have seen, those who reject *Physical Affectability* would not want to maintain this theory of causation in the first place (*Transference* entails *Physical Affectability*, hence the rejection of *Physical Affectability* entails the rejection of *Transference*).

The few remaining mechanistic accounts of causation are wholly compatible with the denial of *Physical Affectability*. Furthermore, those theories of psychophysical causation that deny *Physical Affectability* can provide an explanation of how mental causes have physical effects in the sense that they require. This is because, unlike the energy transference theory of causation, these theories of causation do not analyse the underlying non-causal processes in *physical* terms. Take, for example, Ehring's (1997) account of causation, which identifies the causal process with trope persistence. According to it, complex causal relations involve patterns of partial trope persistence, where these include trope fission and

[29] For a defence of this claim, see Ehring (1997, ch. 1).

trope fusion. The account is mechanistic because causation is analysed in terms of non-causal facts about trope persistence, fission and fusion. However, in order to explain how a cause brings about an effect, one clearly does not need to appeal to the notions of force, energy or momentum; one simply appeals to tropes. It is quite consistent to deny *Physical Affectability* while accepting this theory of causation, for it is certainly not a part of this theory that all trope fission and fusion results in energy and momentum redistribution. And, in answer to the question of how, for example, mental events render physical events non-coincidental the answer is, according to this theory of causation, by trope fission and fusion.[30]

I would therefore suggest that the thought that the denial of *Physical Affectability* is unintelligible stems from the assumption that a theory of causation that analyses causation in terms of underlying non-causal *physical* processes must be correct. With the exception of the energy transference theory of causation, no theory of causation is of this type.

To summarise: I have argued that the plausibility of *Physical Affectability* and *Redistribution* cannot be established by appealing to the energy transference theory of causation and that if these premises are instead to be inferred from facts within physics then it is unclear what these facts are. Finally the remaining objection, that the denial of *Physical Affectability* renders psychophysical causation unintelligible, is only plausible if one assumes the energy transference theory of causation, which is the very theory that those who deny *Physical Affectability* reject. I think that it is therefore fair to conclude that the truth of *Physical Affectability* and *Redistribution* is far from established.

In conclusion, Papineau might have successfully demonstrated that evidence from within current physics suggests that it is highly improbable that any non-physical energy exists. However, he has failed to show why we should move from this claim to the acceptance of *Exclusion* (or the acceptance of *Completeness*). If, as Papineau seems to be assuming, this is to be done via an appeal to the conservation laws, then certain causal claims must be built into his argument. First, the claim that the only way that something non-physical could affect a physical system is by (1) affecting the amount of energy or momentum within it, or (2) redistributing the energy and momentum within it (*Physical Affectability*). Second, the claim that redistribution of energy and momentum cannot be brought about without supplying energy or momentum (*Redistribution*). We cannot appeal to the energy transference theory of causation to establish the plausibility of these claims. And as I hope to have shown, the resulting task of establishing their

[30] Note that if there is any problem with combining Lowe's account with this particular theory of causation (and there may be, as according to Lowe's theory of psychophysical causation, there would be no particular trope or set of physical tropes that a mental trope could be said to fuse with) it certainly does not stem from the fact that Lowe's account rejects *Physical Affectability*.

plausibility, while assuming an account of causation that is less physically biased than that of the energy transference theory, is a truly challenging one. Unless this challenge can be met, Papineau's argument for *Exclusion* can be dismissed, as can all attempts to establish the falsity of interactive dualism via the conservation laws.*

References

Aronson, J. 1971, 'On the Grammar of "Cause"', *Synthese* **22**, pp. 414–430.

Averill, E. and Keating, B. F. 1981, 'Does Interactionism Violate a Law of Classical Physics?' *Mind* **90**, pp. 102–107.

Broad, C. D. 1925, *The Mind and its Place in Nature*, London: Routledge & Kegan Paul.

Cornman, J. W. 1978, 'A Nonreductive Identity Thesis about Mind and Body', in: J. Feinberg, ed., *Reason and Responsibility: Readings in Some Basic Problems of Philosophy*, Encino: Dickenson Publishing Co, pp. 272–283.

Crane, T. 1995, 'Mental Causation', *Proceedings of the Aristotelian Society* Supp. Vol. **69**, pp. 211–236.

Crane, T. 2001, *Elements of Mind*, Oxford: Oxford University Press.

Churchland, P. M. 1994, *Matter and Consciousness*, Cambridge, MA: MIT Press.

Daintith, J. (ed.) 2005, *The Oxford Dictionary of Physics*, Oxford: Oxford University Press.

Dennett, D. C. 1991, *Consciousness Explained*, Boston, MA: Little, Brown and Company.

Dretske, F. 1988, *Explaining Behaviour: Reasons in a World of Causes*, Cambridge, Mass.: MIT Press.

Ehring, D. 1997, *Causation and Persistence: A Theory of Causation*, Oxford: Oxford University Press.

Fair, D. 1979, 'Causation and the Flow of Energy', *Erkenntnis* **14**, pp. 219–250.

Fodor, J. 1994, 'The Mind-Body Problem', in: R. Warner and T. Szubka, eds, *The Mind-Body Problem*, Oxford: Blackwell, pp. 24–40.

Flanagan, O. 1991, *The Science of the Mind*, Cambridge, MA: MIT Press.

Gibb, S. C. forthcoming, 'Non-reductive Physicalism and the Problem of Strong Closure', *American Philosophical Quarterly*.

Goldstein, H., Poole, C. and Safko, J. 2002, *Classical Mechanics*, third edition, San Francisco: Addison-Wesley.

Harbecke, J. 2008, *Mental Causation: Investigating the Mind's Powers in a Natural World*, Heusenstamm bei Frankfurt: Ontos Verlag.

Harré, R. and Madden, E. H. 1975, *Causal Powers: A Theory of Natural Necessity*, Oxford: Blackwell.

Hart, W. D. 1988, *The Engines of the Soul*, Cambridge: Cambridge University Press.

Heil, J. 1998, *Philosophy of Mind*, London: Routledge.

Hopkins, J. 1978, 'Mental States, Natural Kinds and Psychophysical Laws', *Proceedings of the Aristotelian Society* Supp. Vol. **221**, pp. 221–236.

Jackson, F. and Pettit, P. 1988, 'Functionalism and Broad Content', *Mind* **97**, pp. 381–400.

Kim, J. 1992, 'Downward Causation' in: A. Beckermann, H. Flohr and J. Kim, eds, *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*, Berlin: De Gruyter, pp. 119–138.

* This paper was written with support from the Arts and Humanities Research Council's Research Grant AH/F009615/1, 'The New Ontology of the Mental Causation Debate'. Many thanks to Jonathan Lowe, John Heil, James Clarke, Valdi Ingthorsson, three anonymous referees and a member of the *dialectica* Editorial Committee for some very helpful comments on earlier drafts of this paper.

Kim, J. 1993a, 'The Myth of Nonreductive Materialism', in his *Supervenience and Mind: Selected Philosophical Essays*, Cambridge: Cambridge University Press, pp. 265–284.

Kim, J. 1993b, 'The Nonreductivist's Troubles with Mental Causation', in his *Supervenience and Mind: Selected Philosophical Essays*, Cambridge: Cambridge University Press, pp. 336–357.

Kim, J. 2005, *Physicalism, Or Something Near Enough*, Princeton University Press.

Koksvik, O. 2007, 'Conservation of Energy is Relevant to Physicalism', *dialectica* **61**, pp. 573–582.

Levine, J. 2001, *Purple Haze*, New York: Oxford University Press.

Lowe, E. J. 1993, 'The Causal Autonomy of the Mental', *Mind* **102**, pp. 629–644.

Lowe, E. J. 1996, *Subjects of Experience*, Cambridge: Cambridge University Press.

Lowe, E. J. 1999, 'Self, Agency and Mental Causation', *Journal of Consciousness Studies* **6**, pp. 225–239.

Lowe, E. J. 2000, 'Causal Closure Principles and Emergentism', *Philosophy* **75**, pp. 571–585.

Montero, B. 2006, 'What Does the Conservation of Energy Have to Do with Physicalism?' *dialectica* **60**, pp. 383–396.

Papineau, D. 1993, *Philosophical Naturalism*, Oxford: Basil Blackwell.

Papineau, D. 2000, 'The Rise of Physicalism', in: M. W. F. Stone and J. Wolff, eds, *The Proper Ambition of Science*, New York: Routledge, pp. 174–208.

Papineau, D. 2002, *Thinking about Consciousness*, Oxford: Oxford University Press.

Pollock, J. 1990, *How to Build a Person: A Prolegomenon*, Cambridge, MA: MIT/Bradford Books.

Searle, J. R. 1984, *Minds, Brains and Science*, London: BBC.

Searle, J. R. 2004, *Mind: A Brief Introduction*, Oxford: Oxford University Press.

Smith, P. and Jones, O. 1986, *The Philosophy of Mind: An Introduction*, Cambridge: Cambridge University Press.

Taylor, R. 1992, *Metaphysics*, New Jersey: Prentice Hall.

Yablo, S. 1992, 'Mental Causation', *Philosophical Review* **101**, pp. 245–280.