



## Durham Research Online

---

### Deposited in DRO:

04 July 2008

### Version of attached file:

Published Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Lowe, E. J. (2000) 'Causal closure principles and emergentism.', *Philosophy.*, 75 (4). pp. 571-585.

### Further information on publisher's website:

<http://dx.doi.org/10.1017/S003181910000067X>

### Publisher's copyright statement:

The Royal Institute of Philosophy 2000

### Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

---

## Durham Research Online

---

**Deposited in DRO:**  
04 July 2008

**Version of attached file:**  
Published

**Peer-review status of attached file:**  
Peer-reviewed

**Citation for published item:**  
Lowe, E. J. (2000) 'Causal closure principles and emergentism.', *Philosophy*, 75 (4), pp. 571-585.

**Further information on publisher's website:**  
<http://dx.doi.org/10.1017/S003181910000067X>

**Publisher's copyright statement:**  
(c) The Royal Institute of Philosophy 2000

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that :

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full DRO policy for further details.

# Causal Closure Principles and Emergentism

E. J. LOWE

At one time, it was popular to argue against interactionist dualism by appeal to the conservation laws of physics, such as the laws of the conservation of energy and momentum. However, those laws are not sacrosanct and there need be nothing unscientific about questioning them—indeed they have been questioned from time to time by cosmologists and other physicists. Think, for instance, of Bondi and Gold's steady-state model of the universe, which postulated the continuous creation of matter and hence of energy.<sup>1</sup> In any case, appeal to those laws can at best only be used to attack dualist models of psychophysical causation which attribute to the non-physical mind an ability to affect the energy or momentum of a physical system. I say 'at best' because some interactionist dualists, such as W. D. Hart, have postulated the existence of 'psychic energy', which is convertible into physical energy in accordance with the conservation laws.<sup>2</sup> It won't do simply to object that energy is by definition a physical quantity, as this threatens to turn the dispute into a purely verbal one. In view of these and other limitations, it would serve physicalists well if they could frame a more general argument against interactionist dualism, which did not make specific appeal to the conservation laws.<sup>3</sup> Many modern physicalists think that they have an argument of just this sort to hand, in the form of the *causal closure argument*.<sup>4</sup>

In fact, this is not so much a single argument as a family of arguments. A 'causal closure argument', as I shall be interpreting that

<sup>1</sup>See H. Bondi, *Cosmology*, 2nd edn (Cambridge University Press, 1961), Ch. XII.

<sup>2</sup>See W. D. Hart, *The Engines of the Soul* (Cambridge University Press, 1988), Ch. 9.

<sup>3</sup>For further discussion of the (ir)relevance of physical conservation laws to the problem of psychophysical causation, see my *Subjects of Experience* (Cambridge University Press, 1996), pp. 56–63.

<sup>4</sup>A relatively early version of this form of argument may be found in Christopher Peacocke, *Holistic Explanation: Action, Space, Interpretation* (Oxford: Clarendon Press, 1979), pp. 134–43, though Peacocke's version does not exactly conform to the pattern I specify below. I raise an objection to Peacocke's argument in my 'Against an Argument for Token Identity', *Mind* 90 (1981), pp. 120–1.

phrase, has three premises: first, a physical causal closure principle; second, the claim—to which interactionist dualists are themselves committed—that at least some mental events are causes of physical events; and third, the claim that the physical effects of mental causes are not, in general, causally overdetermined.<sup>5</sup> The conclusion of such an argument is that at least some mental events are identical with physical events. Incidentally, I am assuming for the time being that the relata of causal relations are events, broadly conceived to include both processes and states, but not too much turns on this assumption at the moment. Later on, however, I shall introduce a distinction between event-causation and fact-causation which may have a significant bearing on the issues under discussion.<sup>6</sup>

As we shall see shortly, physicalists espousing causal closure arguments do not always agree on the way to formulate their causal closure principles. What is crucial to the success of such an argument is that the causal closure principle appealed to should be neither too strong nor too weak. If the principle is so strong, for instance, as to render redundant the third premise of the argument—the non-overdetermination claim—then the argument will become simply question-begging, because it will amount to an argument from two premises, one of which is accepted by the dualist, to a conclusion which is not accepted by the dualist: and in these circumstances the dualist may justly urge that his opponent is, in effect, merely asserting what he is denying. Moreover, to have any persuasive force, the causal closure principle must be one for which some measure of empirical support can plausibly be mustered, and this places limits on how strong it can be. On the other hand, the causal closure principle must, obviously, not be so weak that the argument in which it is deployed turns out to be invalid. What I shall try to show in this paper is that it is in fact very difficult, if not impossible, for the physicalist to formulate a causal closure principle which is neither too strong nor too weak by these standards. In particular, I shall argue that there are various forms of naturalistic dualism, of an emergentist character, which are perfectly consistent with the

<sup>5</sup> Why should systematic causal overdetermination be ruled out where psychophysical causation is concerned? Perhaps it shouldn't be, and this would be a quick way to dispose of causal closure arguments: see further Eugene Mills, 'Interactionism and Overdetermination', *American Philosophical Quarterly* 33 (1996), pp. 105–17. But I take it that most interactionist dualists would not wish to resort to this strategy if possible, as it looks suspiciously *ad hoc*.

<sup>6</sup> For more on the distinction between event-causation and fact-causation, see Jonathan Bennett, *Events and their Names* (Oxford: Clarendon Press, 1988), pp. 21ff.

strongest physical causal closure principles that can plausibly be advocated. If that is correct, then it means that causal closure arguments against interactionist dualism do not, after all, provide a superior alternative to the discredited arguments which appeal to the conservation laws of physics.

I mentioned a moment ago the variety of causal closure principles appealed to by physicalists. Let me briefly illustrate this variety by some examples. In a recent paper, David Papineau appeals to the following principle, which he calls the principle of ‘the completeness of physics’:<sup>7</sup>

(1A) All physical effects have sufficient physical causes.

But in an earlier book, he appeals instead to a somewhat different principle, which he calls by the same title, namely:<sup>8</sup>

(1B) All physical effects have complete physical causes.

Papineau explains that by ‘complete’ here he means “‘complete’ in the sense that those causes on their own suffice by physical law to fix the chances of those effects’. However, he does also advert to the latter formulation in a footnote to the more recent paper, remarking that ‘a stricter version ... would say that the chances of physical effects are always fixed by sufficient physical causes’.<sup>9</sup>

In another recent contribution to the literature on causal closure, Scott Sturgeon states what he too calls the principle of the completeness of physics as follows:<sup>10</sup>

(1C) Every physical effect has a fully revealing, purely physical history.

However, Sturgeon himself glosses this in terms reminiscent of Papineau’s, as meaning that ‘physical effects have their chances fully determined by physical events alone’. In a paper commenting on Sturgeon’s, Paul Noordhof seizes upon this gloss, remarking that it ‘seems more perspicuous’ than Sturgeon’s official formulation.<sup>11</sup>

<sup>7</sup> David Papineau, ‘Mind the Gap’, in James E. Tomberlin (ed.), *Philosophical Perspectives, 12: Language, Mind and Ontology* (Oxford: Blackwell, 1998), pp. 373–88: see p. 375.

<sup>8</sup> David Papineau, *Philosophical Naturalism* (Oxford: Blackwell, 1993), p. 22.

<sup>9</sup> Papineau, ‘Mind the Gap’, p. 386, n. 4.

<sup>10</sup> Scott Sturgeon, ‘Physicalism and Overdetermination’, *Mind* 107 (1998), pp. 411–32: see p. 413.

<sup>11</sup> Paul Noordhof, ‘The Overdetermination Argument versus the Cause-and-Essence Principle—No Contest’, *Mind* 108 (1999), pp. 367–75: see p. 367.

Accordingly, Noordhof advances the following as his own preferred version of the causal closure principle:

- (1D) Every physical effect has its chance fully determined by physical events alone.

It is worth mentioning, however, that in a further gloss on his version of the principle, Sturgeon asserts that ‘physics does not admit that physical effects have non-physical causes’,<sup>12</sup> and this suggests yet another formulation, namely:

- (1E) No physical effect has a non-physical cause.

Pretty clearly, principles (1A) to (1E) are not all equivalent to one another, whatever one makes of the various glosses supplied by their advocates. And these by no means exhaust the variety of formulations to be found in the literature. One might have hoped for more exactitude and agreement amongst physicalists when it comes to the formulation of a principle so central to their position.

One thing we should notice immediately is that the last version just cited, (1E), is too strong by the standards laid down earlier, because it is bound to render redundant the non-overdetermination premise of any causal closure argument in which it is deployed. To see this clearly, let me now state explicitly what the remaining premises and the conclusion of such an argument are supposed to be. Obviously, these premises and the conclusion are themselves susceptible to some variation in their formulation, but I shall utilize what I hope can be agreed to be uncontentious versions of them. The second premise, maintaining the existence of psychophysical causation, may be stated thus:

- (2) At least some mental events are causes of physical events.

The third premise, excluding the possibility of systematic causal overdetermination, may be expressed as follows:

- (3) The physical effects of mental causes are not, in general, causally overdetermined.

And the anti-dualist conclusion of the argument is supposed to be this:

- (4) At least some mental events are identical with physical events.

However, it is quite evident that (1E) and (2) together entail (4), without the help of (3), so that (1E) can fairly be dismissed by the interactionist dualist as question-begging—as I am sure it would be.

<sup>12</sup> Sturgeon, ‘Physicalism and Overdetermination’, p. 413.

In what follows, I am not going to consider probabilistic versions of the causal closure principle, such as (1D), because they introduce complications which are not relevant to the main thrust of what I have to say. Evidently, the chief reason why probabilistic versions have some favour amongst physicalists is that they do not want it to be objected against their argument that it presupposes a deterministic physics which is at odds with modern quantum theory. But since many dualists and physicalists could agree that quantum-level phenomena are quite probably irrelevant to the problem of psychophysical causation, perhaps we can sidestep the complications raised by probabilistic causation without unduly oversimplifying the debate. It is true, of course, that some interactionist dualists, such as Popper and Eccles, have maintained that quantum-level phenomena are very much relevant to the problem of psychophysical causation, but they are an unrepresentative minority whom we can perhaps afford to ignore for present purposes.<sup>13</sup> And, after all, if interactionist dualism can be rendered consistent with a deterministic version of the causal closure principle that is not excessively strong by the standards laid down earlier, it can hardly be more problematic to render it consistent with a probabilistic version.

With this restriction in mind, let us look next at a version of the causal closure principle which is very close to Papineau's version (1A) above, namely:

(1F) Every physical event which has a cause has a sufficient physical cause.

Here I should say that I understand a *sufficient physical cause* of a given event to be a non-empty set of physical events, each of which is a cause of the given event and all of which jointly causally necessitate the occurrence of the given event. Note that I have included the clause 'which has a cause' in (1F) to allow for the possibility of uncaused physical events, such as, perhaps, a putative 'first' physical event like the so-called 'Big Bang'.

Now, the trouble with (1F) is that it is too weak by the standards laid down earlier, because it does not serve to entail, when conjoined with premises (2) and (3), the anti-dualist conclusion (4). This, in brief, is because (1F) fails to take into account the *transitivity of causation*. To see this, suppose that a certain non-physical mental event *M* is a cause of a physical event *P*, such as a certain bodily movement. But suppose also that *M* itself has a sufficient physical cause. Then, clearly, *P* may have a sufficient physical cause, in accordance

<sup>13</sup> See, e.g., John C. Eccles, *The Evolution of the Brain: Creation of the Self* (London: Routledge, 1989), pp. 187–92.

with principle (1F), namely, a set of physical events which includes the postulated sufficient physical cause of  $M$ . This is perfectly compatible with the non-overdetermination premise, (3), because when a sufficient physical cause of an event  $P$  causes  $P$  by means, *inter alia*, of causing an intervening event  $M$  which is in turn a cause of  $P$ , that sufficient cause and  $M$  do not thereby causally overdetermine  $P$ . More generally, an interactionist dualist who espouses some version of emergentism can happily endorse principle (1F). Such a dualist may consistently maintain that the universe has evolved (without 'outside' influence by supernatural powers) from a condition in which only physical events existed to one in which both physical and mental events exist—the latter conceived as non-physical—while also espousing causal determinism for all events. On such a view, if we trace back the causal history of each mental event, we eventually reach a time at which all of its causal antecedents were wholly physical events, because on this view it is the physical which has ultimately brought the non-physical realm of the mental into being. A little later on, I shall look in more detail at how psychophysical causation might operate in a dualistic world governed by principle (1F).

The physicalist needs, then, a causal closure principle which is stronger than (1F). One likely candidate would be the following, which seems to avoid the problem created by the transitivity of causation:<sup>14</sup>

(1G) At every time at which any physical event has a cause, it has a sufficient physical cause.

With (1G) in place, it might seem, we cannot postulate that a physical event  $P$  has a non-physical cause  $M$  existing at a time  $t_1$  in addition to the sufficient physical cause which, according to (1G), must exist at  $t_1$ , without violating the non-overdetermination premise (3). But in point of fact that is not so, because this neglects the possibility of *simultaneous causation*. Suppose that  $P$  occurs at  $t_2$  and that the postulated sufficient physical cause of  $P$  which exists at  $t_1$  causes  $P$  in part by means of causing  $M$ , which also exists at  $t_1$ . In that case,  $P$  is not causally overdetermined by  $M$  and the physical causes of  $P$  that exist at  $t_1$ . A situation of this sort is illustrated by the following diagram, in which an arrow between events represents the fact that the event at the tail of the arrow is an immediate cause of the event at the head of the arrow:

<sup>14</sup> (1G) is very near to a version of the causal closure principle advanced Jaegwon Kim, namely, 'Any physical event which has a cause at time  $t$  has a physical cause at  $t$ ': see his *Supervenience and Mind: Selected Philosophical Essays* (Cambridge University Press, 1993), p. 280 and, for a slightly different wording, p. 360.

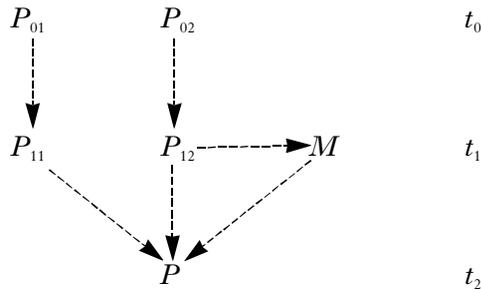


Fig. 1

In the situation depicted in Fig. 1,  $P$  has a sufficient physical cause at time  $t_1$ , namely, the set of physical events  $\{P_{11}, P_{12}\}$ . But  $P$  has as a cause at  $t_1$  the non-physical event  $M$ .  $P$  is not causally overdetermined by  $\{P_{11}, P_{12}\}$  and  $M$ , because it is not the case that in the absence of either one of them  $P$  would still have occurred. In particular, since  $P$  is caused by both  $P_{11}$  and  $P_{12}$ , if  $\{P_{11}, P_{12}\}$  had been absent,  $P$  would not have occurred. Equally, if  $M$  had been absent,  $P$  would not have occurred, for want of a sufficient cause: for we are supposing that  $\{P_{11}, P_{12}\}$  would not have been a sufficient cause of  $P$  if  $P_{12}$  had not caused  $P$  in part by means of  $M$ . Of course, the physicalist may well want to deny the possibility of simultaneous causation, but it should be acknowledged that this is a further substantive claim, without which the causal closure principle (1G) cannot be used to mount an attack on interactionist dualism.

Here I pause to remark that, in his own version of the causal closure argument, Christopher Peacocke considers the possibility of a scenario very similar to the one depicted in Fig. 1, but contends that it is already ruled out by the prohibition on causal overdetermination.<sup>15</sup> This is what he says, replacing his labelling of events by my own:

<sup>15</sup> See Peacocke, *Holistic Explanation*, pp. 136–7. Peacocke’s scenario differs from that of Fig. 1 in that, in his,  $P_{12}$  on its own is a sufficient cause of  $P$ , but this difference is irrelevant for present purposes. It must be observed, however, that Peacocke also makes the following assumption about  $P_{12}$  and  $P$  (again replacing his labelling by mine): ‘[W]e have a complete and wholly physical account of ... the causal route from  $P_{12}$  to  $P$  in neurophysiological terms [and] this account completely explains how the event  $P_{12}$  causes  $P$ ’ (p. 134). Clearly, this assumption is much stronger than anything that can be derived from the causal closure principle (1G). Indeed, on the most natural interpretation of Peacocke’s assumption it simply rules out by definition the sort of situation depicted in Fig. 1, because in the latter there is a *non-physical* causal route from  $P_{12}$  to  $P$  in addition to a physical one. But then one wonders why Peacocke bothered to discuss this sort of situation and object to it on the grounds that it involves overdetermination.

Even if  $P_{12}$  causes  $M$ , and so it is true ... that if  $P_{12}$  had not occurred, then  $M$  would not have occurred, the argument stands. For still there *is* a causal route from  $M$  to  $P$ , one that evidently cannot pass through  $P_{12}$  given the asymmetry of the causation relation; and so overdetermination of  $P$  remains the case. The fact that it is true that if  $P_{12}$  had not occurred, then  $M$  would not have occurred, does not show that there is not overdetermination. We can show this by considering another example. Your pressing a button to light a bonfire electrically may be a signal to me to put a lighted match to the bonfire at the same time; then it is true that if you had not pressed the button, I would not have put the lighted match to the bonfire, but this cannot show that the resulting fire was not overdetermined.

However, it is not enough for Peacocke to establish, by means of such an example, that 'The fact that it is true that if  $P_{12}$  had not occurred, then  $M$  would not have occurred, does not show that there is not overdetermination'. Rather, it is incumbent upon him to establish that in the envisaged psychophysical scenario there *is* overdetermination. Now, the bonfire example obviously *does* involve overdetermination, because your pressing the button would have sufficed to light the bonfire even if it hadn't caused me to light it with a match. But in the situation depicted in Fig. 1, as I am interpreting it,  $\{P_{11}, P_{12}\}$  would *not* have sufficed to cause  $P$  if  $P_{12}$  had not caused  $M$ .

Suppose one were to add to a causal closure argument invoking the causal closure principle (1G) a further premise ruling out simultaneous causation: would (1G) be strong enough to deliver the anti-dualist conclusion (4) in that case? The answer, as we shall see in due course, is that it would not. But before we examine that issue, let us consider whether (1G) is not already too strong by the standards laid down earlier. According to those standards, a causal closure principle should not be so strong that empirical support for it cannot plausibly be mustered. The weaker principle (1F) is certainly not too strong by these standards, since it is empirically plausible to suppose that the universe contained only physical events at an earlier epoch and that latterday mental events are the ultimate effects of causal chains traceable back to events occurring at that earlier epoch. But when we consider that temporal ordering is dense, such that between any two times there is another time, (1G) appears unreasonably strong. Suppose that a physical event  $P$ , such as a bodily movement, is known on good empirical grounds to have a sufficient physical cause in a person's brain at a time 100 milliseconds before the occurrence of  $P$ . And suppose it is further postu-

lated that that sufficient physical cause causes  $P$  in part by causing an intermediate mental event  $M$  which occurs at a time 90 milliseconds before the occurrence of  $P$ . It cannot be illegitimate in principle to postulate something like this, given that temporal ordering is dense, and empirical evidence could quite conceivably support the claim that  $M$  occurred at the postulated time. However, principle (1G) then requires us to suppose that there is a sufficient physical cause of  $P$  existing at a time 90 milliseconds before the occurrence of  $P$ , in advance of any empirical evidence in support of this supposition. That seems extravagant, especially when we consider that (1G) imposes the same requirement with regard to *any* intervening time between 100 and zero milliseconds before the occurrence of  $P$ .

Let me now return to the weaker causal closure principle (1F) and consider how dualistic interaction might operate in a world governed by that principle. What I want to suggest is that, in such a world, mental events might serve to render *non-coincidental* certain physical events which, from a purely physical perspective, might appear to occur by coincidence.<sup>16</sup> For present purposes (and simplifying somewhat for ease of exposition), I take an event to be one which 'occurs by coincidence' if its immediate causes are the ultimate effects of independent causal chains. (One complication which I shall sidestep is the question of what is to count as an 'immediate cause' of a given event, to which different answers might be given according to more or less fine-grained ways of individuating events.) A classic example is this. A man walks to the well from his house one day, following his usual route, but as he passes by a neighbour's house, a gust of wind dislodges a slate which falls on his head, killing him. In this case, the man's death occurred by coincidence, because the causal chain leading to his being where he was when the slate fell was independent of the causal chain leading to the slate's being where it was when the man passed by. We can easily imagine, however, another scenario in which the man's death did not occur by coincidence: for example, a scenario in which a tripwire installed by a jealous neighbour is touched by the approaching man, dislodging the slate so that it falls just as he arrives beneath it.

Changing now to an example involving psychophysical causation, suppose that a physical event  $P$ —such as a bodily movement, or a neural event which is a direct causal antecedent of such a bodily movement—has as its immediate causes two physical events,  $P_{21}$  and  $P_{22}$ , occurring at a time  $t_2$ , and that these physical events are the ulti-

<sup>16</sup> I explore this proposal in more detail and with a slightly different focus in my 'Self, Agency, and Mental Causation', *Journal of Consciousness Studies* 6 (1999), pp. 225–39.

mate effects of two distinct physical causal chains. From a purely physical perspective, these chains might appear to be independent, with the consequence that  $P$  might appear to occur by coincidence. But the truth might be that a mental event  $M$  served to link those two physical causal chains, rendering  $P$  non-coincidental—as is depicted in the following diagram:

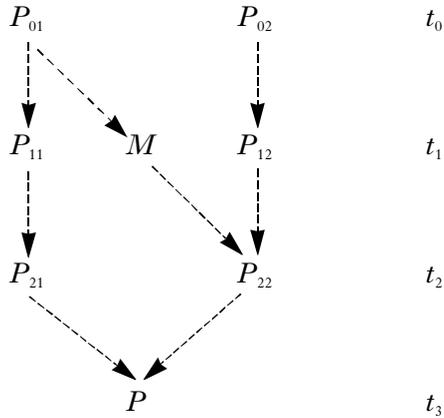


Fig. 2

In the situation depicted in Fig. 2,  $P$  has a sufficient physical cause at times  $t_0$  and  $t_2$ , namely,  $\{P_{01}, P_{02}\}$  and  $\{P_{21}, P_{22}\}$  respectively, but  $P$  does not have a sufficient physical cause at time  $t_1$ . This makes the situation consistent with principle (1F), which requires only that a physical event should have a sufficient physical cause at *some* time at which it has a cause, not at *every* time at which it has a cause—the latter requirement only being imposed by the stronger principle (1G).

In a world containing the situation depicted in Fig. 2, a physical scientist might be able to discover all the physical events depicted there, correctly identify each of them as a cause of  $P$ , and correctly claim that certain combinations of them are sufficient causes of  $P$ . In a sense, he would be in possession of a complete physical explanation of the occurrence of  $P$ , for he would have an explanation which correctly identifies all of  $P$ 's physical causes up to a certain time before the occurrence of  $P$  and a time (in fact, more than one time) at which a subset of those causes is wholly sufficient for  $P$ . And yet, because the mental event  $M$  would be, as it were, invisible to him as a purely physical scientist, his explanation would be, in

another sense, incomplete and would falsely represent the occurrence of *P* as being coincidental. My suggestion is that the world of Fig. 2 may in fact be *our* world. If it *is* our world, physical science can present us with the semblance of a complete explanation of our bodily movements, and yet it will be an explanation which leaves something out, giving our bodily movements the appearance of being coincidental events arising from independent causal chains of events in our brains and nervous systems. But isn't that precisely what current physical science *does* appear to do? As it 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. In short, it leaves us with a kind of 'binding' problem, not unlike the 'binding' problem associated with conscious perceptual experience (the problem of explaining how widely distributed and apparently unrelated neural events in the cortex can support the existence of unified perceptual experiences).<sup>17</sup>

What I want to show next is that the relatively weak physical causal closure principle (1F) is far from being the strongest such principle that is compatible with dualistic interactionism. In fact, dualistic interactionism is even consistent with a causal closure principle that is stronger in some ways than principle (1G), the strongest principle that we have examined so far, apart from the question-begging principle (1E). The principle I have in mind is the following:<sup>18</sup>

(1H) Every physical event contains only other physical events in its transitive causal closure.

By the 'transitive causal closure' of an event *P*, I mean the set of events consisting of the immediate causes of *P*, the immediate caus-

<sup>17</sup> For a description of the more familiar 'binding' problem, see Valerie Gray Hardcastle, 'Psychology's Binding Problem and Possible Neurobiological Solutions', *Journal of Consciousness Studies* 1 (1994), pp. 66–90.

<sup>18</sup> Something like (1H) is suggested by the following recent remark of Jaegwon Kim's: 'One way of stating the principle of physical causal closure is this: If you pick any physical event and trace out its causal ancestry or posterity, that will never take you outside the physical domain. That is, no causal chain will ever cross the boundary between the physical and the nonphysical'. See Jaegwon Kim, *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation* (Cambridge, MA: MIT Press, 1998), p. 40. Of course, another reading of this remark might equate it with the excessively strong principle (1E).

es of those causes, the immediate causes of *those* causes ... and so on: in short, the set which includes every event which stands in the ancestral of the 'immediate cause' relation to  $P$ . The implication of (1H), then, is that the immediate causes of all physical events are always and only other physical events. Principle (1H) clearly prohibits the situations depicted earlier in Fig. 1 and Fig. 2 (remembering that an arrow between events depicted there signifies immediate causation). Indeed, there is a clear sense in which principle (1H) is stronger than principle (1G)—even when the latter is supplemented by a prohibition on simultaneous causation—though it is true that (1H) is weaker in that it does not imply that any physical event has a *sufficient* physical cause. However, in this respect (1H) has an empirical advantage over (1G), inasmuch as it is compatible with the claim, which is plausible in the light of quantum mechanics, that at least some physical causation is irreducibly probabilistic.

How could dualistic interaction be compatible with principle (1H)? For the following reason. It could conceivably be the case that, even though (1H) is true, sometimes a non-physical mental event  $M$  causes it to be the case that certain physical events,  $P_1, P_2, \dots P_n$ , have a certain physical effect,  $P$ .<sup>19</sup> What this would involve is the causation by a mental event of a physical causal fact. So I am assuming now a distinction between event-causation and fact-causation. In a case of fact-causation, what is brought about is not an event, but a fact or state of affairs. Some facts are causal facts, such as the fact that certain physical events,  $P_1, P_2, \dots P_n$ , are causes of another physical event,  $P$ . There is no reason in principle why such a causal fact should not have a causal explanation and, indeed, there may be a positive reason for seeking such an explanation, as I shall explain later. Now, if a mental event  $M$  causes it to be the case that certain physical events,  $P_1, P_2, \dots P_n$ , have a certain physical effect,  $P$ , then, it seems clear,  $M$  is itself a cause of  $P$ —but not an immediate cause of  $P$ , nor an immediate cause of any of the physical events  $P_1, P_2, \dots P_n$ . In fact, it seems clear that  $M$  is not included in  $P$ 's transitive causal closure and hence that  $M$  can be non-physical without violating principle (1H).

There is nothing unintelligible in principle about the notion of a mental event causing a physical causal fact, as the following theological example shows. Suppose that principle (1H) is true in a

<sup>19</sup> For an earlier development of this conception of mental causation, see my *Subjects of Experience*, pp. 64–71 and 82ff. There are some resemblances between my proposal and one made by Fred Dretske: see his 'Mental Events as Structuring Causes of Behavior', in J. Heil and A. R. Mele (eds), *Mental Causation* (Oxford: Clarendon Press, 1993).

world in which every physical event has a sufficient physical cause and no physical causation is either simultaneous or backward. Such a world can have no beginning in time, because it can have no first physical event. (Suppose it did have a first physical event. By hypothesis, that event would have a sufficient physical cause, which would have to be an earlier physical event or conjunction of such events, contradicting the supposition just made.) And yet we could still ask of this world why *it* should exist or be actual rather than any other. One intelligible answer would be to say that this world was actual because God had chosen it to be actual. God's choice, then, would have caused it to be the case that a world containing certain physical causal facts was actual—and this would be mental causation of physical causal facts.

What is envisaged here, it must be emphasised, is not God's causing certain initial physical events, which then go on to cause other physical events in accordance with causal laws chosen by God—for in the envisaged world there are *no* 'initial' physical events, as the world has no beginning in time. Now, of course, it may be asked *when* God's choice was made. If it was made at any time, it would seemingly have to have been made earlier than the occurrence of any physical event. That is not impossible, however, since a physical world which has no beginning in time—no first physical event—need not have a past of infinite duration, any more than a line of finite length need have a first or last point. On the other hand, if God's choice was not made at any time, because He is conceived to be eternal or timeless, then it is true we cannot say that His choice was an *event*—but that in no way detracts from the fact that the example demonstrates the intelligibility of mental causation of physical causal facts.

But how could anything like this model of mental causation be envisaged as applying in cases of ordinary human behaviour? How, for instance, could a mental event, such as a choice of mine, cause it to be the case that certain physical events in my brain cause a certain bodily movement of mine? But what is being requested here? If what is being requested is a demonstration that causation of this kind—the causation of a physical causal fact by a mental event—is intelligible, then we have already provided it by means of the foregoing theological example. If, on the other hand, what is being requested is some account of the 'mechanism' that would be involved in such causation, in the human case, then it may be protested that the request is perhaps unreasonable—first of all because the term 'mechanism' is suggestive of purely physical causation, which is precisely not being envisaged here, and secondly because supplying more detail about the possible nature of such

causation in the human case may not properly be a matter for arm-chair philosophical speculation. My aim has only been to show that dualistic interactionism is compatible with a surprisingly strong form of physical causal closure principle, not to help the would-be interactionist to construct a detailed model of psychophysical causation.

This reply may seem unduly evasive. Is it not an unacceptable implication of the proposed form of mental causation that the mind would have to have the power to affect physical causal laws? Not necessarily, for we need not assume that all causal relations between physical events are subsumable under laws.<sup>20</sup> There may be irreducibly singular causation between physical events, in which case a mental event could cause a physical causal fact without necessarily causing a change in physical laws. But if the real thought behind the objection is that mental causation of the kind envisaged would somehow inevitably conflict with the conservation laws of physics, then I can only say, first, that the burden of proof in this matter lies with my opponent and, second, that if, after all, the physicalist must fall back on an appeal to the conservation laws in an attempt to defeat the dualist, then all the apparent advantages of appealing to causal closure principles have fallen away and we have returned to the tired old debate of earlier times.

But why, it may be asked, should we feel any need to invoke mental causation of the kind just proposed? The answer, once again, might be that only by invoking such mental causation could we 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. At the same time, there might be strong empirical evidence that a *mentale* event was playing precisely such a role, and a strong intuition that the bodily movement in question was *not* an event which occurred ‘by coincidence’. All of this would be perfectly compatible with a naturalistic form of emergentism, moreover, since we could still suppose that every mental event ultimately has a sufficient physical cause. Of course, a physicalist might object that if empirical evidence did confirm that a mental event was playing the kind of causal role now being mooted, we would have to con-

<sup>20</sup> See further G. E. M. Anscombe, ‘Causality and Determination’, in her *Metaphysics and the Philosophy of Mind: Collected Philosophical Papers*, Volume II (Oxford: Blackwell, 1981) and Nancy Cartwright, *Nature’s Capacities and their Measurement* (Oxford: Clarendon Press, 1989).

clude that this event was identical with some physical event, almost certainly some neural event in the brain. But *why* would we have to conclude this, in the absence of any direct empirical evidence for the existence of such a physical event? What I have been trying to show in this paper is that causal closure arguments do not provide the kind of justification for such a claim that physicalists may fondly think they do.<sup>21</sup>

*University of Durham*

<sup>21</sup> I am grateful for comments received when an earlier version of this paper was delivered at the School of Advanced Study in the University of London.