

4

Is Non-reductive Physicalism Viable within a Causal Powers Metaphysic?

*Timothy O'Connor and John Ross Churchill**

Throughout the 1990s, Jaegwon Kim developed a line of argument that what purport to be *non-reductive* forms of physicalism are ultimately untenable, since they cannot accommodate the causal efficacy of mental states. His argument has received a great deal of discussion, much of it critical. We believe that, while the argument needs some refinement, its basic thrust is sound. In what follows, we will lay out our preferred version of the argument and highlight its essential dependence on a causal powers metaphysics, a dependence that Kim does not acknowledge in his official presentations of the argument. We then discuss a recent physicalist strategy for preserving the causal efficacy of the mental in the face of this sort of challenge, a strategy that endorses a causal powers metaphysics of properties while offering a distinctive account of the physical realization of mental properties. We argue that the resulting picture cannot be satisfactorily worked out, and that seeing why it fails strongly suggests that non-reductive physicalism and a causal powers metaphysics are not compatible, as our original argument contends.

1. A CAUSAL POWERS ONTOLOGY

Let us first explain what we mean by the term 'causal powers'. One way of using this term involves no definite commitments on the metaphysics of causation. A person using the term this way might say, for example, that a defoliant has the causal power to kill plants, where this claim is neutral as to whether (a) the

* Versions of this paper were read at Queen's University Belfast, St Louis University, the Catholic University of Milan, and the University of Durham. We thank the audiences on these occasions for helpful criticism and advice, especially Tim Crane, Carl Gillett, John Heil, Stephan Leuenberger, E. J. Lowe, Cynthia Macdonald, Graham Macdonald, Peter Menzies, Paul Noordhof, David Papineau, and Jessica Wilson.

presence of the defoliant would or might stand in an ontologically basic relation of producing or bringing about the death of a plant, (b) there is a law of nature that relates properties of the defoliant and plant death, (c) plants regularly die after being sprayed with the defoliant, (d) there are subjunctive conditionals relating the properties of the defoliant and the death of plants in a certain way, (e) citing the presence of the defoliant satisfactorily explains the occurrence of plant deaths in certain contexts, or some (f) fitting a still different analysis of causation.

We do not use the term in this neutral manner. Our usage corresponds to the first of these: a power to produce or to bring about some event, where this is assumed to be a real relation irreducible to more basic features of the world. Our favoured technical term for this is 'causal oomph'. So understood, causation is not amenable to analysis in non-causal terms, but instead involves the exercise of ontologically primitive causal *powers* or *capacities* of particulars. Powers are either identical to, or figure into the identity conditions of, certain of the object's properties, which are immanent to those things as non-mereological parts. (Whether one thinks of these as immanent universals or tropes is not crucial in this context.)

It bears emphasis that this view is *not* committed to assuming that all causation must amount to something like pushing/pulling or the exertion of a force. What is assumed, rather, is solely this: when an instance of a property—the event of the particular's having the property—is a cause, the world unfolds in a certain way after the instance of that property, and that property instance is one of the factors that jointly *make* the world unfold this way. This is just another way of saying what's come before, that the property instance and others jointly produce or bring about certain effects; they jointly oomph the world into going on in *this* way rather than *that*. Because of this, there are certain counterfactuals true of the world ('were the property not to have been instanced, such-and-such effects would not have occurred'). But these counterfactuals are derivative from, and not to be equated with, or seen as the basis of, the causal facts themselves: it is because the property instance was among the factors that jointly produced the relevant happenings that certain corresponding counterfactuals are true. Causally efficacious properties have the power to make the world unfold in ways that otherwise it would not, and this is a fundamental feature about these properties upon which all else (counterfactuals true of them, regularities and patterns that encompass them, explanations that cite them) is derivative.

There is much debate, and not a little confusion, over how to delineate the finer points of this general picture. While we cannot delve deeply into these matters, we make the following two remarks to forestall confusion that might infect understanding of our subsequent argument.

First, there is a pervasive manner of speaking that appears on the surface to say that *objects* have and exercise causal powers. (Witness our example above with respect to defoliant.) In our view, such talk should be construed by the causal powers metaphysician as a shorthand way of expressing the claims that:

- (1) the object's having the property *is* its having the causal power;
- (2) the event of the property's being had by the object over a particular interval of time and in appropriate circumstances will causally contribute to the effect, where it occurs; and
- (3) the exercise of the causal power just is this causal contribution.¹

Second, a single property may contribute to a very wide array of effects, depending on the context in which it is instanced. A particle's being negatively charged may contribute to its accelerating at varying rates away from a similarly charged nearby particle, accelerating towards an oppositely charged nearby particle, even accelerating towards a similarly charged particle (though at a slower rate than would occur were the particle not to have been so charged), and countless other manifestations, all depending on the context of its occurrence. But in ordinary speech, again, there is a tendency to talk of a corresponding array of causal powers being exercised, 'each' of which is identified through the effect actually manifested. This sort of speech has encouraged some metaphysicians to posit a multiplicity of properties, or worse, to posit a distinct type of entity (a causal power), any number of which are 'conferred by' a single property. We should resist such moves on grounds of parsimony, and here science is a much better guide to property/power identifications. The key is to understand a basic power or disposition not in terms of this or that salient manifestation, but rather in terms of a unitary causal influence, something that is constant across circumstances while its manifestations will vary.²

In considering the prospects for a non-reductive physicalist view of the mental, we are assuming, rather than arguing for, this causal powers metaphysics. We are investigating its implications for the question at hand. Can the (by our lights) right-thinking metaphysician who has seen his way clear to this view of causation make out a non-reductive physicalist view on which mental states are causally efficacious in this sense? We will try to persuade you that the prospects are bleak.

¹ One *might* hold to a philosophical view leading one to insist that in certain cases, it is indeed the object that exercises the power, and not the event of the object's having the property/causal power. Such is the claim of the agent causationist, e.g., with respect to the forming of a free decision. But this is a substantive and controversial thesis, not a spelling out for one sort of case what is common to every case of causation. (For a discussion of the relationship of agent causation to the more usual 'event causation' within a causal powers metaphysics, see O'Connor [2008].)

² We are influenced here by Corry (2002, 2008) and Heil (2003).

2. CAUSAL POWERS AND THE DILEMMA OF REDUCTION OR CAUSAL EXCLUSION OF THE MENTAL

We will now present our preferred version of a Kim-style argument.³ We begin with three related premises concerning causation and properties:

- (1) Causation is a real relation irreducible to more basic features of the world (*causal non-reductionism*).
- (2) Causation involves the exercise of ontologically basic causal powers or capacities of particulars (*production account of causation*).
- (3) Properties are individuated in terms of causal powers, such that there are no distinct properties that confer exactly the same causal profile (*causal theory of properties*).

The next four premises flow from the distinctive commitments of non-reductive physicalists:

- (4) No mental property is identical to any physical property (*distinctness thesis*).
- (5) Mental properties supervene on physical properties (*supervenience thesis*).

The hoary slogan, of course, is ‘no mental difference without a physical difference’, intended to capture an appropriate dependence relation. What exact form the supervenience relation should take in this context, however, is a difficult and controverted issue. We will follow Kim in supposing that complication arising, e.g., from mental content externalism can be safely ignored. If this is correct, we may assume for the sake of argument that mental properties ‘strongly supervene’ on the physical properties of the individual (or on the physical properties and relations of the individual’s parts). Next we have:

- (6) Mental properties are realized by physical properties: a particular event *M* of a person *S*’s having mental property *M* is either ‘constituted by’ (a kind of ontological posteriority) or is identical to various physical particulars—possibly including portions of the person’s environment—having certain physical properties and standing in certain physical relations (*realization thesis*).

We will be non-committal on whether the realization of mental properties by physical properties involves constitution or identity of the corresponding events, since non-reductive physicalists’ pronouncements on this matter are varied and often obscure.⁴ Finally, physicalists typically wish to assert:

³ See Kim (1993, 1997, 1998, 1999, 2003, and 2005).

⁴ See for example the variation among Fodor (1974); Pereboom and Kornblith (1991); Pereboom (2002); Shoemaker (2001, 2007); and Gillett (2002).

- (7) For every *physical* event, its objective chance of occurring is fully fixed by physical events (*causal completeness of physics*).

According to (7), nothing non-physical is *required* in order to causally account for the occurrence of any physical event.

We now contend that (1)–(7) are inconsistent with supposing:

- (8) There is a causally efficacious mental event, *M*, that is the instancing of a particular mental property, *M*. The causal activity of *M* is distinct from the activity of the physical event, *P*, that is the instancing of *M*'s realizer property (or properties), and this activity in one way or another impinges the realm of physical events⁵ (*assumption for reductio*).

Premise (8) is an instance of the general claim that the causal efficacy of mental properties does not reduce in every case to the causal efficacy of some physical properties. The singular causal action of the mental event of *M*'s being instanced does not reduce to the singular causal action of some physical event or events, such as the instancing of the physical property *P* that realizes *M* in the circumstances.

The argument that (8) is inconsistent with (1)–(7) proceeds as follows:

- (9) The instance of *M* either
(a) directly produces a subsequent mental event, *M**, or
(b) it directly produces a wholly physical event, *P**.

The realization thesis (6) and non-reductive-productive account of causation (1–2) together strongly suggest that option (a) is a non-starter. On this view, mental events are ontologically dependent on their subvening realizers, wholly constituted by (if not identical to) them, and this is no less true of mental *effects* as of mental *causes*. Bringing about such a mental event *eo ipso* involves causally affecting the physical event which realizes it. So

- (10) Not (9a).

But the thesis of causal completeness (7) implies that:

- (11) If 9b, then the physical event *P** is overdetermined by *M* and some other physical event.

Now, if we accept the non-reductive-productive account of causation, it will seem passing strange to suppose that, in regular fashion, there are physical events that are systematically 'overoomphed' by distinct events, even if—indeed, *especially* if—these causes might stand in a supervenience relation. If, say, a physical event *P*, the realizer of the mental event *M*, produces or oomphs *P**, what causal work

⁵ We will, for the sake of convenience, continue to refer only to *P*, the single realizer of *M*, though it should be understood that on some accounts of realization *M* may be realized by multiple properties ('the *P*'s', say) each time it is instanced. Gillett (2002) is one such account.

is left over for M? Note that on reductive accounts of causation, on which causal facts are not something additional to the totality of non-causal facts, the situation looks very different. Suppose, for example, that our effect P* is counterfactually dependent on both P and M. If we accept something like the counterfactual analysis of causation, there is nothing strange or objectionable about deeming M, as well as P, to be a cause of P*. For in doing so we are not making a commitment to anything additional—M's status as a cause of P* falls out of the facts that we already accept, along with our analysis. It comes for free. By contrast, on the non-reductive-productive account, we would be positing an additional fundamental relation between M and P*, when doing so is entirely unnecessary for accounting causally for P*.⁶ Thus, we should conclude that:

- (12) There is not systematic mental-physical overdetermination, as the consequent of (11) implies.

But this is the end of the road. We are forced to conclude, therefore, that:

- (13) M does not make a distinctive contribution to occurrences in the physical world, whether wholly physical or supervening mental occurrences (*completing reductio of (8)*).

Finally, the causal theory of properties (premise 3) both rules out an epiphenomenalist retreat and suggests the proper ultimate conclusion: we ought either to reductively *identify* M with P or *deny* that M is a bona fide property—one that earns its causal keep—in the first place.

Here is our argument laid out in compact form:

- (1) Causation is a real relation irreducible to more basic features of the world (*causal non-reductionism*).
- (2) Causation involves the exercise of ontologically basic causal powers or capacities of particulars (*production account of causation*).
- (3) Properties are individuated in terms of causal powers, such that there are no distinct properties that confer exactly the same causal profile (*causal theory of properties*).
- (4) No mental property is identical to any physical property (*distinctness thesis*).
- (5) Mental properties supervene on physical properties (*supervenience thesis*).
- (6) Mental properties are realized by physical properties: a particular event M of a person S's having mental property M is either 'constituted by' (a kind of ontological posteriority) or is identical to various physical particulars—possibly including portions of the person's environment—having certain physical properties and standing in certain physical relations (*realization thesis*).
- (7) For every *physical* event, its objective chance of occurring is fully fixed by physical events (*causal completeness of physics*).

⁶ These points are made clearly and effectively by Loewer (2001), a review of Kim (1998).

- (8) There is a causally efficacious mental event, M, that is the instancing of a particular mental property, M. The causal activity of M is distinct from the activity of the physical event, P, that is the instancing of M's realizer property (or properties), and this activity in one way or another impinges the realm of physical events (*assumption for reductio*).
- (9) The instance of M either
 - (a) directly produces a subsequent mental event, M*, or
 - (b) it directly produces a wholly physical event, P*.
- (10) Not (9a) (*from 1, 2, 6*).
- (11) If 9b, then the physical event P* is overdetermined by M and some other physical event (*from 7*).
- (12) There is not systematic mental–physical overdetermination, as the consequent of (11) implies (*from 1,2*).
- (13) M does not make a distinctive contribution to occurrences in the physical world, whether wholly physical or supervening mental occurrences (*completing reductio of (8)*).

The argument just presented, like earlier relatives, seeks a reductionist or eliminativist conclusion by way of arguing for the *exclusion* of irreducibly mental causation. Yet it does this by explicitly invoking the thesis of causal powers realistically construed. So let us refer to it hereafter as the *power exclusion argument*.

The commitments that drive the power exclusion argument are tenets of the causal powers metaphysics, on the one hand, and non-reductive physicalism, on the other. If we wish to preserve a realist and non-reductive view of the mind and its causal influence, we must reject one or another tenet of these two packages. Or so we believe. Sydney Shoemaker, however, disagrees, and he has recently attempted to provide a way out for the non-reductive physicalist who is a realist with respect to causal powers. Since Shoemaker has bona fides as both a causal powers metaphysician and as a physicalist and has attempted to work out the metaphysics of realization and causation with much greater care than is usual in these discussions, it is fitting that we investigate his approach in detail.

3. SHOEMAKER ON NON-REDUCTIVE MENTAL CAUSATION

Shoemaker thinks that the key to vindicating the causal efficacy of mental properties without reduction lies in a distinctive account of the *realization* of mental properties by physical properties. In broad strokes, his proposal is that mental and other realized properties—those that do real causal/explanatory work—belong to a special class of *disjunctive* properties, with their disjuncts as their realizers: the relation of realizer to realized is simply the relation of disjunct

to disjunction.⁷ On this view, realized properties have a proper subset of each of their realizers' *forward-looking causal features*—what instances of the properties can causally suffice for—while having a superset of their realizer properties' *backward-looking causal features*—what can causally suffice for instances of the properties. Shoemaker then exploits the conclusion that realized properties have a subset of their realizers' powers to argue that mental causation is not reducible to causation by the physical realizers, owing to a certain proportionality thesis explained below concerning what counts as a cause of what. The following schema captures Shoemaker's picture.

B1 → C1	B1 → (C1 ∨ C2)	C1 → E1	C1 → (E1 ∨ E2)
B2 → C2	B2 → (C1 ∨ C2)	C2 → E2	C2 → (E1 ∨ E2)
(B1 ∨ B2) → (C1 ∨ C2)		(C1 ∨ C2) → (E1 ∨ E2)	

Figure 4.1. Bold font (**C1**) indicates a property, while regular font (C1) indicates a property instance, or the event of an object's having the property at a particular time. The '→' denotes causal sufficiency. C1 and C2 represent instances of different realizing physical properties of (C1 ∨ C2), the multiply realized property instance. B-type events are possible causal determinants of C-type events, while E-type events are possibly determined by C-type events, according to the patterns indicated. A realized event such as (C1 ∨ C2) has more possible determinants than any of its realizers while it suffices for fewer effects.

How is accepting this picture of realization supposed to make things easier for non-reductive physicalism? Says Shoemaker: begin by observing that if the realized property has a subset of the forward-looking causal features of the realizer, then the realizer property instance is causally sufficient for everything the realized property instance is causally sufficient for, *plus more*. So, for example, C1 is causally sufficient for (E1 ∨ E2), just as (C1 ∨ C2) is, but unlike the latter it is also sufficient for an instance of E1. Now, if C1 and (C1 ∨ C2) overlap in this way in what they causally suffice for, and if causal considerations ought to drive our conclusions about the identity of properties, a natural conclusion is that (C1 ∨ C2) is a proper *part* of C1 (in the sense of entailment). More generally: the instances of realized properties are parts of the instances of the corresponding realizers and so are not identical to them.⁸

From here, Shoemaker invokes a version of Stephen Yablo's 'proportionality' constraint⁹ on what we ought to count as the cause in a causal interaction: while it is true that C1 is causally sufficient for (E1 ∨ E2), (C1 ∨ C2) is, Yablo and Shoemaker say, a better candidate for being the cause. For (C1 ∨ C2) is also causally sufficient for the specified effect, but only just so—it causally suffices for the effect *and nothing more besides*. The only features of C1 that contribute

⁷ See Shoemaker 2007: section II of ch. 2, especially 17–18, 55–6, and section V of ch. 4, especially 79 and 82.

⁸ See Shoemaker (2007: 13, 53).

⁹ Shoemaker (2001: 81). See Yablo (1992).

to the bringing about of (E1 v E2) are features had by (C1 v C2), a part of C1. Compare, says Shoemaker, a more familiar sort of case: Jones fires a single shot as part of the volley of a firing squad, but his shot arrives just ahead of the others, killing the condemned. In the realization case as Shoemaker describes it, just as with the firing squad analogy, we are invited to conclude that while the whole (C1; the firing squad's firing) was causally sufficient for the effect ((E1 v E2); the death of the condemned), proportionality constraints argue in favour of counting a particular part of the event ((C1 v C2); Jones's firing) rather than the whole as *the cause*.¹⁰ This is how realized events in general qualify as causes in certain scenarios. They do not overdetermine their effects alongside their realizers. Instead, they always cause effects which their realizers suffice for but do not cause.

We now have before us Shoemaker's account of realization and the way it provides for non-reductive mental causation. But how exactly, we may wonder, does the account underwrite a response to the power exclusion argument? Notice, first, that there is no rejection of the distinctness, supervenience, or realization theses (premises 4–6). What of (7), the causal completeness principle? Shoemaker insists on the importance of a controversial distinction between causal determination or sufficiency and causation proper. But this position is consistent with the way that we have formulated (7), and it seems clear that Shoemaker accepts it.

In our view, the *way* that Shoemaker applies the (mere) sufficiency/causation distinction to cases of realization is untenable. We argue for a disjunctive conclusion: either Shoemaker is soft-pedaling the irreducible efficacy of the mental (as embodied in premise 8) or his commitment to the causal powers metaphysics (as expressed in premises 1–3) is less than it appears.

To bring the problem into focus, consider first that, for all his distinctive claims, Shoemaker clearly gives ontological priority to the physical realizer event. He tells us that P realizes M just in case P is metaphysically sufficient for (but not identical to) M and 'constitutively makes it real' (2007: 4, 10). He goes so far as to say that realized states are 'nothing over and above' their realizers (2007: 2). If all this is so, then how is a case of M's causing an effect, E, not also a case whereby P, M's constituting realizer, is likewise causing E? Indeed, how is this not a case where P is causally prior to M, so that, by the power exclusion argument, we should conclude that P is the sole true cause?¹¹ We see two alternatives:

(A) What he terms causal 'determination' or 'sufficiency' is, at least in the context of realizer/realized events, metaphysically primary. His talk of

¹⁰ See Shoemaker (2001: 81, 2007: 13–14, 52–3).

¹¹ A bolstering consideration comes from certain indeterministic scenarios. We take it to be evident that, *assuming the causal completeness of physics*, the chance of E given M cannot be greater than the chance of E given a total physical cause (here, our P). But there seems to be no reason to think that it cannot be less. Now consider a case where $\text{Pr}(E/M)$ is significantly less than $\text{Pr}(E/P)$. Surely, in such a case, where E in fact occurs, it is highly implausible to insist that nevertheless M, not P, is the cause of E. While this is a special case, if our conclusion about it is accepted, it seems to indicate that there is something wrong about Shoemaker's method for assigning causes.

'causation' in these contexts should be interpreted as something like 'causal relevance'—i.e., an *explanatory* relation. But if this is what he intends, then mental causation is not *metaphysically* irreducible, only explanatorily ineliminable for certain purposes. This sort of view has its defenders, and we shall say something about it below, when considering alternatives to non-reductive physicalism proper.

- (B) If Shoemaker insists that mental causation as he describes it is no less metaphysical than basic physical causation, then we suspect some sort of retreat from a causal powers metaphysics, for only in this way could you say that P is somehow 'merely' causally sufficient whereas M is the proper cause. If P is ontologically prior to M, *able* to bring about E, and in the circumstances necessary to do so, how can it get out-oomphed by M?

We might try out a modified picture that gives up some of what Shoemaker claims in exchange for better prospects for a distinctive causal efficacy of the mental. For example, we might ignore Shoemaker's talk of P's being ontologically prior to and constitutively making real M and focus instead on his notion that M is a part of P, owing to the subset-of-powers thesis and the causal theory of properties. (In this reconstruction, we would emphasize Shoemaker's invocation of the firing squad analogy (2007: 53) and also such statements as 'It is only because the c-fiber [stimulation] instance realizer contains the pain instance realizer that it has the relevant effects' (2007: 48).) In the resulting picture, we would have what amounts to a radical inversion of the reductionist's vision, such that it is the *physical* properties that resolve into an assemblage of mental properties plus some non-mental causal features. That would allow for mental causation that is irreducible in one sense, at least: it does not reduce to causation by the corresponding *macrophysical* events.

We think it would be very hard to make the picture a plausible one, though we set that worry aside. It is enough to note that making the view out will probably require us to analyse the associated macrophysical property as a structural property, the instancing of which just consists in, or is constituted by, the instancing of properties of the object's parts and relations between them. The mental property then comes out, on the view being considered, as an overlapping structural property, perhaps somehow abstracted from the full physical structural property. But if we do so, *both* mental properties and the larger structural physical properties in which they are embedded turn out to be derivative structures, entities that are constructions out of *microphysical* properties and relations. The spectre of reductionism menaces again.

Now, Shoemaker allows that macro-level properties are in a sense structural, but he resists their reductive *identification* with microphysical states of affairs. He suggests that instances of macro-level properties are microphysically realized, where this latter realization relation, like that of same-level realization, involves only constitution, not identity. However, his case for this rests on two claims

about property identity that should be unacceptable to a causal powers theorist (2007: 48–9).

First, he lays down that, in general, a property instance has just one constituent object and one constituent property, so a mental property instance cannot be identical to a state of affairs involving many distinct properties and objects. He seems to put this forward as a definitional truth or platitude. But a causal powers theorist does not take quasi-grammatical considerations to be final arbiters concerning the structure of reality. One might just as well take Shoemaker's supposed platitude together with facts (assumed for now) about microphysical constitution and draw the conclusion that there are not, strictly speaking, mental properties at all—not in the sense of entities that contribute directly to how the world unfolds.

Shoemaker's second claim is that the *modal* properties of macro-level property instances and their microphysical realizers will generally differ. (Consider familiar claims made in discussions of the statue of Goliath.) This claim rests on intuitive judgements about possible variation in the material constitution of composite objects. But the status of composite objects, no less than that of 'their' properties, is very much in question on the powers metaphysics. One cannot simply assume that there are robustly *objective* modal facts about them and use these to ward off what otherwise appears to be a powerful reductionist challenge. For the causal powers theorist, the question of which candidates for being true (non-conventional) macro-level objects will turn on whether they manifest properties that play an ineliminable causal role. Shoemaker is alive to this issue, drawing a distinction between 'genuine' and 'phoney' properties (2007 ch. 4, sections 5–6) on the basis of what is 'directly detectable' by us. (We have a genuine macro-level property when we are able to reliably and directly detect its instantiation independently of its realizer.) Shoemaker freely acknowledges, however, that he has no argument that direct detectability suffices for genuineness. Nor does he say anything to allay the immediate worry that this will make the concept of genuineness to be epistemic, not metaphysical. Once again, we think he is forced to retreat to a defence that does not sit well with a metaphysics that revolves around causal powers.

It is time for a recap of our argument thus far. We have defended a power exclusion argument for the untenability of non-reductive physicalism. It is a variant of Kim's argument that makes explicit an assumption of a causal theory of properties. We then tried to show that Sydney Shoemaker's recent attempt to harmonize the two positions fails. Note that the causal theory of properties pretty directly entails a sparse, rather than abundant ontology of properties, as most predicates do not correspond to anything that makes a causal difference in the world. Shoemaker's strategy, in effect, is to try to make room for a less-than-austere, though still limited, inventory of properties, one that allows for irreducible macrophysical and mental properties, while being consistent with the physicalist's vision that everything supervenes on the microphysical. We believe

that this strategy is bound to fail, since the causal theory requires that irreducible properties earn their keep, and there is no room for this at the macrophysical level if physicalism is true.

4. ALTERNATIVES TO NON-REDUCTIVE PHYSICALISM

The sparse metaphysics of causal powers forces a choice limited to three stances concerning macroscopic structures: reduction or elimination, if strict physicalism is maintained, or a rejection of one or more of the characteristic claims of physicalism—that is, towards the acceptance of an ontological variety of emergence.

Non-reductive physicalists see an obstacle to the first option, reductionism, in the fact that, as functional properties, intentional properties are multiply realized. What counts as a belief that *Q* in humans may be quite distinct, at any physical level of description, from what counts as that same belief in, say, an intelligent extraterrestrial or a sophisticated artificial machine built out of steel and silicon. Kim and some other reductionists recommend that we seek local, species-specific reductive identities for intentional properties—*human* belief that such-and-such as identical with physical property so-and-so—and so preserve the status of these intentional properties as causal powers. That is, we characterize both *M* and *P* in terms of *highly* specific mental and physical types, respectively, and move to a type-type identity theory.

The second, eliminativist option is to interpret apparent reference to mental properties as properly denoting mental *concepts* only. There are far fewer properties possessed by an object than the vast number of concepts it falls under. Genuine *properties* are immanent to their instances and make a non-redundant difference to how the objects act in at least some circumstances. As critics of Kim have observed, causal exclusion arguments appear to generalize beyond mental properties to all properties posited in the special sciences (sciences other than basic physics).¹² And since, contra Kim, it is highly plausible that special science categories are not typically ontologically reducible (owing in part to their own multiple realizability),¹³ the argument ultimately leads (they say) to an eliminativist conclusion. This is often taken as a *reductio ad absurdum*: surely the terms of well-established biological and chemical theory pick out genuinely efficacious properties!

Now one response to this proposed *reductio* of the exclusion argument is to note the availability of a third alternative. Rejecting premise (6) and (7), the realization and causal completeness theses, suffices to block the final conclusion

¹² For discussion, see Baker (1993); Burge (1993); van Gulick (1993); Kim (1996, 1997, 1999, 2003, 2005); Block (2003); Ross and Spurrett (2004).

¹³ See Fodor (1974); Dupré (1993); and Rosenberg (1994).

of the power exclusion argument. We will discuss shortly the viability of this strategy in relation to mental properties. We will not discuss whether this is viable as a general strategy for special science properties, though we observe that recent philosophy of science has seen a significant challenge to the completeness thesis in particular.¹⁴

Suppose that one takes the case for the completeness of physics with respect to some or all of the special sciences to be convincing, setting aside sciences impinging on mentality. In that case, contra Fodor and many others, it would not be absurd to embrace eliminativism. For so-called high level theories can be enormously useful and illuminating, and even necessary to the progress of human knowledge of how the world works, without answering to ontological 'levels' or layers populated by distinctive properties and their objects.¹⁵ And the further fact that such theories are not generally reducible to more fundamental theories is a highly interesting one about our world (and necessary for science to get off the ground, as in practice we inevitably work our way in, not out), but it cuts no ontological ice. An alternative to the levels picture of physical reality has already been hinted at above: there is a vast array of microphysical entities (for simplicity, 'the particles') bearing primitive, dynamical features and standing in primitive relations. Talk of composite objects and their properties, at least in the general case, is the imposition of a conceptual scheme that selectively picks out coarse-grained patterns running through the vast storm of particles. These concepts really are (*objectively*) satisfied by the world, but not in virtue of a one-one relation between general concepts and properties, or individual concepts and particulars.

This second, eliminativist response to the powers exclusion argument might be thought to entail, implausibly, the devaluation of the special sciences. Such a conclusion would be too hasty, however. For it is simply false that science is of value only as a source of representing the world's causal joints in more and more accurate ways. It is, in addition, a source of means for intervening in and manipulating the world so as to change it for the better, and much of its value is due to this rather than to its representational fruits. We value science—we fund it, prioritize it, give special social status to many of its practitioners, etc.—because of its role in improving the world, and not just because of its role in representing the world. (The development of methods for effectively preventing and treating myriad diseases serves as just one example of such improvement.) But qua sources of improvement, some of the special sciences are at least as valuable, and perhaps more so, than fundamental physics. For very often we are *better* able to intervene and manipulate in ways that improve the world by using the resources of the non-fundamental special sciences.

¹⁴ See Cartwright (1999) and Dupré (1993, 2001). And for a powerful challenge to the case for completeness in the special scientific domain of chemistry in which it is widely thought to be most secure, see Hendry (2006).

¹⁵ On this point, see Heil (2003, chs 2–7).

So much for the first two alternatives left open by the power exclusion argument. Though we have suggested that eliminativism, in particular, is a viable approach to the predicates that appear in many of the special sciences, we believe that, when it comes to mentality, both reduction and elimination are implausible.

To secure a robust efficacy for mental properties, we should reject not only causal completeness but also mental-physical realization. Such are the negative commitments of what we call *ontological emergence*. (Carl Gillett's so-called *strong emergence* rejects completeness only. We have argued elsewhere that this position is too weak to secure a causal role for putative emergent properties.)

The term 'emergence' is used to cover a multitude of sympathies (in some cases, sins). So we want to indicate in clear, albeit very abstract, terms what an emergentist picture would look like, in our way of thinking.

Properties are *ontologically emergent* just in case:

- (i) They are ontologically basic properties (token-distinct from, and unrealized by, any structural properties of the system).
- (ii) As basic properties, they constitute new powers in the systems that have them, powers that non-redundantly contribute to the system's collective causal power, which is otherwise determined by the aggregations of, and relations between, the properties of the system's microphysical parts. Such non-redundant causal power necessarily means a difference even at the microphysical level of the system's unfolding behaviour. (This is compatible with the thesis that the laws of particle physics are *applicable* to such systems. It requires only that such laws be supplemented to account for the interaction of large-scale properties with the properties of small-scale systems.)

In respects (i) and (ii), emergent properties are no less basic ontologically than unit negative charge is taken to be by current physics. However, emergent and microphysical properties differ in that

- (iii) emergent properties appear in and only in organized complex systems of an empirically specifiable sort and persist if and only if the system maintains the requisite organized complexity. The sort of complexity at issue can be expected to be insensitive to continuous small-scale dynamical changes at the microphysical level.¹⁶

We are inclined to further suppose that

- (iv) the appearance of emergent properties is *causally originated and sustained* by the joint efficacy of the qualities and relations of some of the system's

¹⁶ Concepts of emergence have a long history—one need only consider Aristotle's notion of irreducible substantial forms. Their coherence is also a matter of controversy. For an attempt to sort out the different ideas that have carried this label, see O'Connor and Wong (2002). And for a detailed exposition and defence of the notion we rely on in the text, see O'Connor and Wong (2005).

fundamental parts. (This would involve fundamental properties having latent dispositions to contribute to effects, dispositions that are triggered only in organized complexes of the requisite sort.)

One cannot give uncontroversial examples of emergent properties, of course. Though there are ever so many macroscopic phenomena that seem to be governed by principles of organization highly insensitive to microphysical dynamics, it remains an open question whether such behaviour is nonetheless wholly determined, in the final analysis, by ordinary particle dynamics of microphysical structures in and around the system in question.¹⁷ Given the intractable difficulties of trying to compute values for the extremely large number of particles in any medium-sized system (as well as the compounding error of innumerable applications of approximation techniques used even in measuring small-scale systems), it may well forever be impossible in practice to attempt to directly test for the presence or absence of a truly (ontologically) emergent feature in a macroscopic system. Furthermore, it is difficult to try to spell out in any detail the impact of such a property using a realistic (even if hypothetical) example, since plausible candidates (e.g., phase state transitions or superconductivity in solid state physics, protein functionality in biology, animal consciousness) would likely involve the simultaneous emergence of multiple, interacting properties. Suffice it to say that if, for example, a particular protein molecule were to have emergent properties, then the unfolding dynamics of that molecule *at a microscopic level* would diverge in specifiable ways from what an ideal particle physicist (lacking computational and precision limitations) would expect by extrapolating from a complete understanding of the dynamics of small-scale particle systems. The nature and degree of divergence would provide a basis for capturing the distinctive contribution of the emergent features of the molecule.

Now, many contemporary philosophers seem to think that such a view is too extreme to be plausible. When pressed, such critics often cite the alleged consequence that an emergentist view compromises *the unity of nature*. But unity does not require the reductionist vision of the world as merely a vast network binding together local microphysical facts, with a pervasive and uniform causal continuity underlying all complex systems. It is enough that at every juncture introducing some new kind of causally discontinuous behaviour, there is a causal *source* for that discontinuity in the network of dispositions that underlie it. In short: unity in the order of the unfolding natural world need not involve causal continuity of behaviour, only continuity of dispositional structure.¹⁸ For

¹⁷ For numerous examples of such phenomena, see Laughlin et al. (2000).

¹⁸ This is not to concede that it is *ipso facto* a theoretical virtue for a metaphysics that it entails greater unity in nature, nor that it is *ipso facto* a theoretical vice if the converse is true. The issue of the unity of nature, and the related issue of unity in science, is deep and complex. Our point in the text is that there is *a* kind of unity in nature if the emergentist account I have proposed is correct. For more on the topics of unity in science and nature, see Cat (2007).

the emergentist, the *seeds* of every emergent property and the behaviour it manifests are found within the world's fundamental elements, in the form of latent dispositions awaiting only the right context for manifestation.

We make no assertion one way or the other as to whether anything is like this for any chemical or biological properties, though we note that present evidence allows for the possibility that some perfectly respectable biological and chemical features are ontologically emergent in this way.

We do, however, propose that the conscious intentional and phenomenal aspects of the mind strongly favour an emergentist account. A human person's experiences and other conscious mental states exhibit features quite unlike those of physical objects, whether as revealed in ordinary sense perception or as uncovered in the physical and biological sciences. And the maximally direct nature of our first-person awareness of the intentional and phenomenal features of our conscious states blocks the a posteriori ascription to them of underlying physical microstructure hidden to introspection. The upshot of this familiar reflection, if it stands, is that our experiences and other conscious mental states have fundamentally distinctive characteristics. But these very characteristics are also *prima facie* causally efficacious. (Indeed, on a causal powers metaphysics, to countenance them as properties is to accept them as efficacious.) Thus, certain mental properties appear to be (1) resistant to analysis in terms of physical structural properties and so plausibly ontologically basic, (2) causally efficacious, and (3) borne only by highly organized and complex systems. Though we cannot argue the matter at length here, we find extant materialist attempts to overcome this *prima facie* case to be implausible.¹⁹ (It goes without saying that we take the grounds for an emergentist account of the mental to be defeasible.)

Some philosophers acknowledge that the sort of broadly 'Cartesian' picture sketched above captures how we naively think about conscious experience but contend that it is an illusion. For our part, we think that such philosophers underestimate the difficulties for a theory of empirical knowledge that maintains that we are subject to a radical and pervasive cognitive illusion at the very source of all of our empirical evidence. And if the central argument of this paper is correct, then for any of these philosophers likewise committed to a causal powers metaphysics, the seemingly *paradoxical* position of denying the causal efficacy of mental states must be added to those difficulties.

REFERENCES

- Baker, L. R. 1993. 'Metaphysics and Mental Causation'. In J. Heil and A. Mele (eds), *Mental Causation*. Oxford: Clarendon Press, 75–95.
- Block, N. 2003. 'Do Causal Powers Drain Away?'. *Philosophy and Phenomenological Research* 67: 133–50.

¹⁹ For argument on this point, see O'Connor and Kimble (n.d.).

- Burge, T. 1993. 'Mind-Body Causation and Explanatory Practice'. In J. Heil and A. Mele (eds), *Mental Causation*. Oxford: Clarendon Press, 97–120.
- Cartwright, N. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- 1989. *Nature's Capacities and Their Measurement*. Oxford: Clarendon Press.
- 1999. *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.
- Clapp, L. 2001. 'Disjunctive Properties: Multiple Realizations'. *Journal of Philosophy* 98: 111–36.
- Corry, R. 2002. 'A Causal-Structural Theory of Empirical Knowledge'. PhD thesis, Indiana University.
- 2008. 'Scientific Analysis and Causal Influence'. In Toby Handfield (ed.), *Dispositions and Causes*. Oxford: Oxford University Press.
- Crisp, T. and Warfield, T. 2001. 'Kim's Master Argument: a Critical Notice of *Mind in a Physical World*'. *Noûs*: 304–16.
- Dupré, J. 1993. *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge, MA: Harvard University Press.
- (2001). *Human Nature and the Limits of Science*. Oxford: Oxford University Press.
- Fodor, J. 1974. 'Special Sciences'. *Synthese* 28: 97–115.
- Gillett, C. 2002. 'The Dimensions of Realization: a Critique of the Standard View'. *Analysis* 62: 316–23.
- Heil, J. 2003. *From an Ontological Point of View*. Oxford: Oxford University Press.
- Hendry, R. 2006. 'Is there Downward Causation in Chemistry?'. In D. Baird, L. McIntyre, and E. Scerri (eds), *Philosophy of Chemistry: Synthesis of a New Discipline*. Boston Studies in the Philosophy of Science, vol. 242. Springer, 173–89.
- Kim, J. 1993. *Supervenience and Mind*. Cambridge, MA: MIT Press.
- 1997. 'Does the Problem of Mental Causation Generalize?'. *Proceedings of the Aristotelian Society* 97: 281–97.
- 1998. *Mind in a Physical World*. Cambridge: Cambridge University Press.
- 1999. 'Supervenient Properties and Micro-based Properties: a Reply to Noordhof'. *Proceedings of the Aristotelian Society* 99: 115–18.
- 2003. 'Blocking Causal Drainage and Other Maintenance Chores with Mental Causation'. *Philosophy and Phenomenological Research* 67: 151–76.
- 2005. *Physicalism, or Something Near Enough*. Princeton, NJ: Princeton University Press.
- Laughlin, R. B., Pines, D., Schmalian, J., Stojkovic, B., and Wolynes, P. 2000. 'The Middle Way'. *Proceedings of the National Academy of Sciences* 97: 32–7.
- Lewis, D. 1980. 'Mad Pain and Martian Pain'. In N. Block (ed.), *Readings in the Philosophy of Psychology*, vol. 1. Cambridge, MA: Harvard University Press, 216–22.
- Loewer, B. 2001. 'Review of J. Kim, *Mind in a Physical World*'. *Journal of Philosophy* 98: 315–24.
- O'Connor, T. 2008. 'Agent-Causal Power'. In Toby Handfield (ed.), *Dispositions and Causes*. Oxford: Oxford University Press.
- and Kimble, K. n.d. 'The Argument from Consciousness Revisited.' Unpublished manuscript.
- and Wong, H. Y. 2002. 'Emergent Properties'. *Stanford Online Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/properties-emergent/>.

- O'Connor, T. and Wong, H. Y. 'The Metaphysics of Emergence'. *Noûs* 39 (2005): 659–79.
- Pereboom, D. 2002. 'Robust Non-Reductive Materialism'. *Journal of Philosophy* 99: 499–531.
- and Kornblith, H. 1991. 'The Metaphysics of Irreducibility'. *Philosophical Studies* 63: 125–45.
- Rosenberg, A. 1994. *Instrumental Biology or the Disunity of Science*. Chicago, IL: University of Chicago Press.
- Ross, D. and Spurrett, D. 2004. 'What to Say to a Skeptical Metaphysician: a Defense Manual for Cognitive and Behavioral Scientists'. *Behavioral and Brain Sciences* 27: 603–47.
- Shoemaker, S. 1980. 'Causality and Properties'. In P. Van Inwagen (ed.), *Time and Cause*. Dordrecht: Reidel, 109–35.
- 1998. 'Causal and Metaphysical Necessity'. *Pacific Philosophical Quarterly* 79: 59–77.
- 2001. 'Realization and Mental Causation'. In C. Gillett and B. Loewer (eds), *Physicalism and its Discontents*. Cambridge: Cambridge University Press.
- 2002. 'Kim on Emergence'. *Philosophical Studies* 108: 53–63.
- 2007. *Physical Realization*. Oxford: Oxford University Press.
- Tooley, M. 1987. *Causation: A Realist Approach*. Oxford: Oxford University Press.
- 1993. 'Causation: Reductionism versus Realism'. In E. Sosa and M. Tooley (eds), *Causation*. Oxford: Oxford University Press, 172–92.
- van Gulick, R. 1992. 'Non-Reductive Materialism and Inter-theoretic Constraint'. In A. Beckermann, H. Flohr, and J. Kim (eds), *Emergence or Reduction? Essays on the Prospects of Non-Reductive Physicalism*. Berlin: De Gruyter, 157–79.
- 1993. 'Who's in Charge Here? And Who's Doing All the Work?'. In J. Heil and A. Mele (eds), *Mental Causation*. Oxford: Clarendon Press.
- Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.
- Yablo, S. 1992. 'Mental Causation'. *The Philosophical Review* 101: 245–80.