AN ARGUMENT FOR POWER INHERITANCE

UMUT BAYSAN[1]

Abstract: Non-reductive physicalism is commonly understood as the view that mental properties are *realized* by physical properties. Here, I argue that the realization relation in question is a power inheritance relation: if a property *P* realizes a property *Q*, then the causal powers of *Q* are a subset of the causal powers of *P*. Whereas others have motivated this claim by appealing to its theoretical benefits, I argue that it is in fact entailed by two theses: (i) realization is a same-subject necessitation relation; (ii) properties have their causal powers derivatively on the causal powers of their bearers. Although the power inheritance claim that is defended here has many opponents, I take it that the two theses that entail it are either plausible or widely assumed.

I

It is commonplace to take *non-reductive physicalism* (NRP) to be a *realization* thesis. According to this understanding, NRP is the view that mental properties are not identical with, but are realized by, physical properties, where realization is a relation whose exact formulation is up for discussion.[2] Here, I shall argue that the following principle is true for the realization relation: *if* a property *P* realizes a property *Q*, *then* the causal powers of *Q* are a subset of the causal powers of *P*.[3] Let us call this principle *Power Inheritance* (PI). Jessica Wilson (1999; 2011; 2015), Lenny Clapp (2001) and Sydney Shoemaker (2001; 2007)[4] argue for PI on the grounds that it has certain theoretical benefits. Each provides a distinctive way of explaining how PI is meant to show that NRP doesn't imply that mental properties are causally excluded by their physical realizers. Moreover, Wilson holds that PI ensures the physicalist contention that 'mental properties are "nothing over and above" their base properties' (1999: 41). In what follows, I provide a new argument for PI which doesn't appeal to its alleged theoretical benefits, thanks to which I am hoping to sidestep the issues related to the 'nothing over and above' locution and the success of the proposed solutions to

[1] Philosophy, University of Glasgow; emin.baysan@glasgow.ac.uk

[2] Such realization-based formulations of physicalism can be found in Boyd (1980) Poland (1994), Wilson (1999), Melnyk (2003) and Shoemaker (2007). See Baysan (2015) for a review of some of these formulations.

[3] This subset condition is not a sufficient condition for realization. There could be pairs of properties which satisfy the subset condition but are not related via realization. As noted by Wilson (1999) and Shoemaker (2001), a conjunctive property (*P* & *Q*) would have *P*'s causal powers as a subset, but in general, conjunctive properties are not realizers of their conjuncts.

[4] All of these philosophers argue that the said subset condition is in fact a *proper* subset condition. Despite my agreement, here I shall only argue for PI, leaving it open whether the subset relation is a proper or an improper one. An ancestor of PI is Kim's *causal inheritance principle*, which suggests that 'if mental property *M* is realized in a system at t in virtue of physical realization base *P*, the causal powers of this instance of *M* are identical with the causal powers of *P*' (1992: 18). Kim uses this principle in an argument against NRP, aiming to show that the principle leads to a reductive view, not a non-reductive one. So it is plausible to interpret Kim to be making an *improper* subset claim.

the causal exclusion problem.[5] Admittedly, PI is controversial; many have rejected it for reasons independently of the issues that are related to the causal exclusion debate (e.g. Noordhof 1997; 1999; 2013; Gillett 2002; 2003; 2010; Pereboom 2002; 2011; Menzies & List 2010). Interestingly, despite its controversial status, PI is in fact a consequence of two claims which are either plausible or widely assumed. Or so I shall argue.

II

The argument that I shall present for PI is motivated by physicalism's commitment to the supervenience of the mental on the physical. Much ink has been spilled over identifying the right variety of supervenience that physicalism should be committed to. One proposal is that a variety of *strong* supervenience in Kim's (1984) sense is a better candidate than other varieties of supervenience. If this is right, then such supervenience entails that (i) one cannot have a mental property without thereby having *some* physical property, and (ii) a mental property *M* is *necessitated* by any physical property *P* that *M* supervenes on. The definition of strong supervenience is silent about the modal strength of necessitation. There is a plausible argument that physicalists should appeal to *metaphysical* necessitation, not merely *nomological* necessitation, as otherwise their views wouldn't be distinguishable from some anti-physicalist views (Noordhof 2003).[6] Despite my agreement, I will not be committed to this latter claim about metaphysical necessitation and will stick to the following conditional: if physicalism is true, then mental properties are *at least* nomologically necessitated by physical properties. Given that realization is the relation that relates mental properties to physical properties according to NRP, the following is a very plausible claim:

(1) If a property *P* realizes a property *Q*, then, as a matter of nomological necessity, all bearers of *P* are also bearers of *Q*.[7]

Let's now introduce a putative case of realization:

(2) The property of having *C-fibre stimulation* realizes the property of having *pain*.[8]

[5] An anonymous referee points out that there are implausible consequences of PI regarding mental causation. The worry is that PI (conjoined with NRP) would imply that mental causation is not distinguishable from physical causation. This is a worry that Wilson, Clapp, and Shoemaker have responded in their defences of what is sometimes called the "subset strategy" to solve the exclusion problem. It is disputed whether such solutions are successful, and I take it that the advantage of my defence of the PI is that it doesn't appeal to any accounts of mental causation. It might as well turn out that the view actually implies what the anonymous referee is worried about, but then the source of the worry could easily be NRP, not PI.

[6] Some examples of such anti-physicalist views are that of Chalmers (1996) and strong emergentism, where some mental properties supervene on physical properties only nomologically, but not metaphysically.

[7] This is an understanding of realization that Gillett (2002; 2003; 2010) explicitly disagrees with. Gillett thinks that we should take realization to be akin to a mereological relation where a realized property and its realizers are instantiated at different mereological levels. On his account, typically, the instantiation of a realizer property in object *O* does not necessitate (in any sense) the instantiation of a realized property in *O*. It is not surprising then that Gillett also happens to disagree with PI, for the simple reason that parts and wholes have different causal powers.

From (1) and (2), it follows that

(3) as a matter of nomological necessity, all bearers of *C-fibre stimulation* are also bearers of *pain*.

A few words regarding (3) are in order. Due to massive oversimplifications, (3) is obviously false. As Shoemaker (1981) illustrates, *C-fibre stimulation* in a Petri dish will not bring about an instantiation of *pain*. So, a Petri dish, or the stuff in the dish, can be a bearer of *C-fibre stimulation* without thereby being a bearer of *pain*, even if physicalism is true.[9] It is known that such difficulties can be avoided by including some background conditions in the realizer so that neural firings in Petri dishes wouldn't count as appropriate realizers of mental properties. After all, the realizing properties can be extrinsic properties, as it can also be acknowledged due to the observation that mental content can be 'wide'. If one wishes to formulate physicalism without reference to such extrinsic properties, it is possible to do so in terms of a 'global' supervenience thesis. However, the point that I am making here is easier to express in terms of 'local' supervenience, so I shall allow realizer properties to include such extrinsic properties, if needed.

The illustration of the next step for the argument that I am presenting here needs some scene-setting. The question that is relevant here is this. What do we mean when we attribute causal powers to properties? Consider a variation of an example from Shoemaker (1980). *Being knife-shaped* has the power to cut bread---conditionally on being instantiated with certain other properties, of course. When we attribute this power to the property of *being knife-shaped*, do we really mean that *the property itself* has this power? Unless we want to *identify* properties with bundles of causal powers, I don't think that we have any good reason to give an affirmative answer to this question. *Properties don't cut bread*. Their bearers might. To generalise, properties don't (literally or fundamentally) have causal powers; their bearers do.

If one chooses to analyse causal power attributions to properties in terms of causal power attributions to their bearers, then it might appear that the next question to answer is: In virtue of what do objects have causal powers? In other words, what are the truthmakers of true dispositional statements? When I truly say that a glass is fragile, what makes this fragility-ascription true? This is a notoriously tricky issue to settle. Luckily, the argument that I am presenting for PI doesn't depend on how to settle it. Whatever the correct account of dispositional attributions to *objects* is, we can take that account and then derive the correct account of causal power attributions to *properties*.

If we are to explain the causal powers of properties in terms of the causal powers of their bearers, then we ought to find a *systematic* way of doing so. One option would be to hold the following thesis:

---

[8] Henceforth, I shall omit using the clause 'the property of having' before property names, and use italics for property names.

[9] Shoemaker makes this observation regarding metaphysical necessity, not nomological necessity, but the point applies to both.

(4a) A property *P* has, among its causal powers, a causal power *C* just in case all bearers of *P* have *C*.

Although (4a) is a relatively good approximation, there are some problems with it. One counterexample will suffice. Suppose that all yellow things in a world *w* are destroyed, except one mustard bottle. (Assume that mustard is not yellow in *w*.) If (4a) were true, it would have been the case that, in *w*, the property of *being yellow* has the causal power to squirt mustard. But this is surely very odd.[10] One way of getting round this problem would be by means of introducing a modal element in the formulation:

(4b) A property *P* has, among its causal powers, a causal power *C* just in case, *necessarily*, all bearers of *P* have *C*.

Some philosophers will find the right-hand side of (4b) too strong---if the necessity in question is to be understood as metaphysical necessity. Particularly, there is the worry that (4b) might commit us to what is sometimes called *dispositionalism*, namely the view that properties have their causal profiles as a matter of metaphysical necessity, which is meant to suggest that a property confers on its bearers the same causal powers in all possible worlds.[11] Whether dispositionalism is true or not is disputable, and the argument for PI shouldn't be dependent on this.[12] One obvious fix would be to relativise the strength of the modal operator to nomologically possible worlds.

(4c) A property *P* has, among its causal powers, a causal power *C* just in case, *as a matter of nomological necessity*, all bearers of *P* have *C*.

The latter provides a systematic way of explaining how a given property can be assigned a causal profile. Moreover, it has certain advantages over other alternatives: there are no obvious counterexamples to it; it is not committed to metaphysically loaded theses; and, more importantly, it is compromising enough to accommodate metaphysically loaded theses. For example, if one holds that nomologically possible worlds and metaphysically possible worlds are co-extensional, then she can still hold (4c). In that case, (4b) and (4c) would be equivalent, but then there wouldn't be any problem with being committed to dispositionalism, as dispositionalists typically reject the nomological-metaphysical distinction, due to their belief that laws of nature hold non-contingently.[13]

---

[10] I am grateful to David Bain for this colourful example.

[11] See Shoemaker (1980) and Swoyer (1982).

[12] Other accounts of PI differ on PI's relation to dispositionalism. Despite his attempts to dissociate his view of realization from dispositionalism, Shoemaker's (2001; 2007) account of realization is very closely linked to this view. Clapp's (2001) account of realization is, by his own admission, an extension of dispositionalism. Wilson (2011) explicitly states that her account of realization is compatible with any theory about properties and their relations to causal powers, insofar as these theories accept that what an object can do is somehow related to what properties it has.

[13] Menzies & List (2010), who disagree with the PI, wouldn't be happy with (4c), or any of its variations I have considered. Their disagreement with the PI is motivated by cases like the following. Suppose that I have a desire to drink water and hence move my arm towards the water bottle. My desire *M* causes my behaviour *E*. Let's say that *M* is realized by a neural property *N*. If PI is true, then the causal power which

It might be objected that (4c) rules out the nomological possibility of epiphenomenal properties. An epiphenomenal property, by definition, doesn't have any causal powers, but it seems to be possible that all bearers of an epiphenomenal property may have some causal powers in common---due to sharing some other properties. So, one qualification that we might want to add to (4c) would be to restrict it to only non-epiphenomenal properties. This qualification should be unproblematic for the intended conclusion, as the argument that I am presenting is for a thesis about properties *with* causal powers.

Now, consider the following platitude: pain experiences dispose their subjects to cry. Some pain subjects might not manifest their dispositions to cry for whatever reason---they might be superspartans---but being disposed to cry is the right sort of disposition to associate with pain experiences. Hence,

(5) *pain* has, among its causal powers, the causal power to cry.

Now, from (4c) and (5), it follows that

(6) as a matter of nomological necessity, all bearers of *pain* have the causal power to cry.

Now, remember from (3) that all bearers of *C-fibre stimulation* are, as a matter of nomological necessity, bearers of *pain*. Then, (3) and (6) yield the following.

(7) As a matter of nomological necessity, all bearers of *C-fibre stimulation* have the causal power to cry.

Recall (4c), which links causal powers of properties to causal powers of their bearers. From (4c) and (7), we obtain:

(8) *C-fibre stimulation* has, among its causal powers, the causal power to cry.[14]

What does (1)-(8) show? It shows that PI is true. Let me explain why. We see that an arbitrarily chosen causal power of *pain* turns out to be also a causal power of *C-fibre*

---

is responsible for the occurrence of *E* should be a causal power of both *M* and *N*. But Menzies and List suggest that this is not the case because *N* is not a legitimate, proportional, cause of *E*. Therefore, PI is false. Or so they argue. That *N* is not a legitimate cause of *E* falls out of their understanding of causation as a difference-making relation. Given that they have a difference-making approach to causation, it isn't very surprising that they wouldn't be on board with the suggestions I have made regarding (4c) above. However, as an anonymous referee points out, their objection is unsuccessful on other grounds too. It appears that they conflate the question of whether *N* is a legitimate (proportional) cause of *B* with the question of whether *N* confers on its bearers a causal power whose manifestations can be events like *B*. After all, *N* and *M* might confer the same given causal power on their bearers, yet only one's doing so could be counted as an instance of proportional causation.

[14] This conclusion shouldn't be conflated with the following statement: *C-fibre stimulation* confers the causal power to cry on all bearers of *pain*. This latter statement is false, as there can be bearers of *pain* which aren't bearers of *C-fibre stimulation*. One person who disagrees with PI is Noordhof (1997; 1999; 2013), whose objection turns out to target this latter statement, which I take to be false. He argues that *C-fibre stimulation* doesn't have all causal powers of *pain*. His reason is that whereas pain confers some powers on, say, robots, *C-fibre stimulation* doesn't. What (8) in my argument says is that the causal power to cry is conferred on all bearers of *C-fibre stimulation*, not that *C-fibre stimulation* confers this power on objects which are not its bearers.

stimulation. But note that the properties *C-fibre stimulation* and *pain* are arbitrarily chosen too, apart from the fact that they are related via the realization relation. From what has been argued for, it follows that, for any two properties *P* and *Q*, if *P* is a realizer of *Q*, then any causal power that is attributed to *Q* should also be attributed to *P*. In other words, the causal powers of *Q* are a subset of the causal powers of *P*.

III

The argument I have just presented shows that PI is a consequence of two theses: (i) realization is a same-subject necessitation relation with at least nomological strength; (ii) the causal powers of a property should be understood in terms of the causal powers of its bearers. Those who question PI[15] are kindly invited to consider which of these two claims they want to dispense with.[16]

References

Baysan, U. (2015). 'Realization Relations in Metaphysics', *Minds and Machines* 25/3: 247–60.

Boyd, R. (1980). 'Materialism Without Reductionism: What Physicalism Does Not Entail', in N. Block (ed.) *Readings in the Philosophy of Psychology*. Cambridge: Harvard University Press.

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

Clapp, L. (2001). 'Disjunctive Properties', *Journal of Philosophy* 98: 111–36.

Gillett, C. (2002). 'The Dimensions of Realization: A Critique of the Standard View', *Analysis* 62: 316–23.

Gillett, C. (2003). 'The Metaphysics of Realization, Multiple Realizability and the Special Sciences', *Journal of Philosophy* 100: 591–603.

Gillett, C. (2010). 'Moving Beyond the Subset Model of Realization: The Problem of Qualitative Distinctness in the Metaphysics of Science', *Synthese* 177: 165–92.

[15] Pereboom (2002; 2011) explicitly rejects PI. He takes that the causal powers of a realized property and the causal powers of its realizer are related with a *constitution* relation: if a property *P* realizes a property *Q*, then the causal powers of *Q* are constituted by the causal powers of *P*. Since constitution doesn't entail identity, he suggests, no causal power of *Q* can be identical with any causal power of *P*. But this is not really an *argument* against PI. Pereboom *assumes* from the outset that the relationship between the causal powers of *Q* and *P* is constitution, which is an *asymmetric* relation.

[16] I am grateful to Jonas Christensen, Joaquim Giannotti, Carl Gillett, Stephan Leuenberger, Fiona Macpherson, Neil McDonnell, Paul Noordhof, Eric Saidel, Jessica Wilson, and two anonymous referees of this paper for their comments on various versions of this paper. The work for this article was carried out thanks to a grant provided by The Durham Emergence Project funded by the John Templeton Foundation (grant number: 40485)

Kim, J. (1984). ‘Concepts of Supervenience’, *Philosophy and Phenomenological Research* 45: 153–76.

Kim, J. (1992). ‘Multiple Realization and the Metaphysics of Reduction’, *Philosophy and Phenomenological Research* 52: 1–26.

Menzies, P. and List, C. (2010). ‘The Causal Autonomy of the Special Sciences;, in C. Mcdonald and G. Mcdonald (eds.) *Emergence in Mind.* Oxford: Oxford University Press.

Melnyk, A. (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism*. New York, NY: Cambridge University Press.

Noordhof, P. (1997). ‘Making the Change, the Functionalist's Way’, *The British Journal for the Philosophy of Science* 48: 233–50.

Noordhof, P. (1999). ‘Micro-based Properties and the Supervenience Argument: a Response to Kim’, *Proceedings of the Aristotelian Society* 99: 109–14.

Noordhof, P. (2003). ‘Not old ... but not that new either: Explicability, Emergence and the Characterisation of Materialism’, in S. Walter & H. Heckman (eds.) *Physicalism and Mental Causation: The Metaphysics of Mind and Action*, 85–108. Exeter, Imprint Academic.

Noordhof, P. (2013). ‘Mental Causation: Ontology and Patterns of Variation’, in S. Gibb, E. J. Lowe & R. D. Ingthorsson (eds.) *Mental Causation and Ontology,* 88–125. Oxford, Oxford University Press.

Pereboom, D. (2002). ‘Robust Nonreductive Materialism’, *Journal of Philosophy* 99: 499–531.

Pereboom, D. (2011). *Consciousness and the Prospects of Physicalism*. New York, NY: Oxford University Press.

Shoemaker, S. (1980). ‘Causality and Properties’, in P. van Inwagen (ed.), *Time and Cause: Essays Presented to Richard Taylor,* 109–35. Dordrecht: Reidel.

Shoemaker, S. (1981). ‘Some Varieties of Functionalism’, *Philosophical Topics,* 12: 93–119.

Shoemaker, S. (2001). ‘Realization and Mental Causation’, in C. Gillett & B. Loewer (eds.) *Physicalism and its Discontents,* 74–98. Cambridge: Cambridge University Press.

Shoemaker, S. (2007). *Physical Realization*. Oxford: Oxford University Press.

Poland, J. (1994). *Physicalism: The Empirical Foundations*. Oxford: Oxford University Press.

Wilson, J. (1999). ‘How Superduper Does a Physicalist Supervenience Need to Be?’, *The Philosophical Quarterly* 49: 33–52.

Wilson, J. (2011). ‘Non-Reductive Realization and the Powers-Based Subset Strategy’, *The Monist* 94: 121–54.

Wilson, J. (2015). ‘Metaphysical Emergence: Weak and Strong’, in T. Bigaj & C. Wuthrich (eds.) *Metaphysics in Contemporary Physics: Poznan Studies in the Philosophy of the Sciences and the Humanities*.