

Supervenience, Properties and Mental Causation

Anthony Dardis
Philosophy, Hofstra University
Anthony.Dardis@hofstra.edu

January 11, 2008

Abstract

Suppose mental property P supervenes on and is not identical with physical property Q . This paper defends non-reductive physicalism, by showing how P could be causally relevant to properties of effects of instances of P .

The supervenience argument

(Kim, 2005) argues that if mental properties are irreducible to (and not identical with)¹ physical properties (*Irreducibility*), then epiphenomenalism follows: no instance of a mental property ever causes any instance of any property. The conclusion is secured with the help of three premises:

Supervenience. Mental properties strongly supervene on physical/biological properties. That is, if any system s instantiates a mental property M at t , there necessarily exists a physical property P such that s instantiates P at t , and necessarily anything instantiating P at any time instantiates M at that time. (33)

Closure. If a physical event has a cause that occurs at t , it has a physical cause that occurs at t (43)

Exclusion. No single event can have more than one sufficient cause occurring at any given time—unless it is a genuine case of causal overdetermination. (42)

¹Kim takes it that if property P is not reducible to property Q then they are not identical. This paper is concerned with an argument to be made in the case that P supervenes on and is not identical with Q , and so this paper is entirely silent on the question what kinds of things stand in reducibility relations (properties, theories, languages ...).

Here's the part of the argument I'll focus on. Suppose thinking of cheesecake causes me to form the intention to reach for that cheesecake, and the intention causes my arm to move. By *Supervenience*, these mental properties supervene on physical properties. These physical properties cause the effects, if anything does. Assume *Irreducibility*, so the mental cause is not identical with the physical cause. Hence for each of these effects, there appear to be two causes. But this is not a genuine case of causal overdetermination. By *Exclusion* and *Closure* the mental cause is eliminated.

Kim responds by denying *Irreducibility*. Here, I deny that the mental property is fully distinct from the physical property, and hence that Kim's argument fails; I go on to propose an account of property causal relevance that will permit both supervenient and subvenient properties to be causally relevant.

Properties, non-identity and distinctness

Kim has had relatively little to say about what properties *are*. This matters: some accounts of properties underwrite his argument, others do not. Here is one that does not (Lewis, 1986, 1983; Jubien, 1989; Taylor, 1993).

To every set of actual and/or possible entities there corresponds a property.

For convenience we can think of properties as sets.

Given this conception of properties, Kim's concern about *Irreducibility* pales somewhat. Whether mental properties are *reducible* to physical properties is a complex question. Whether they are *identical with* physical properties looks like a simpler question: for non-identity, it suffices that there is at least one non-physical instance of each mental property. This appears to be logically possible, since it is not contradictory to assert that there are non-physical minds.

Smart gave what he called "irreducibly psychic properties" (Smart, 1959, 148) a bad name by associating them with qualitative characteristics of experiences and by implying that such things would be distinct from brain processes or indeed anything physical at all. On the

present conception, a property is not physical when it has possible instances that are not physical things. The actual instantiation of such properties does not entail that their actual instances are not completely and exhaustively made of purely physical matter. Nor do such non-physical properties have to be “nomological danglers.”

A mental property may not be identical with any physical property, *and also* not be fully distinct from any physical property. The mental property of wanting cheesecake—the set—includes all the actual, physical, instances of that desire, plus some non-physical instances as well. Since the physical property is a proper subset of the mental property, the mental property is neither identical with, nor fully distinct from, the physical property. (See (Sanford, 2005) on why distinctness and non-identity are different in general, not just for the case of properties.)

Above I commented in passing that I don’t believe that the mental cause and the physical cause genuinely overdetermine the effect; the reason is that if they are not distinct, then they aren’t genuinely overdetermining causes. Moreover, if mental properties are not identical with, but also are not distinct from, physical properties, then Kim’s *Exclusion* does not apply. Where an event has both a mental cause and a physical cause, and they are not identical but also not distinct, it’s not so that the effect has *more than one* sufficient cause.

So I think that, given this conception of properties, Kim’s argument fails. That doesn’t save mental causation, however. Kim is relying on a powerful intuition: by *Closure*, there’s always a physical cause for anything that happens. Isn’t that enough? If there were anything else, wouldn’t it be redundant? So the problem is to work out a positive account of how the mental cause can do any work—how the non-reductive physicalist may accept causation at many levels ((Kim, 2005, 63-4), referring to Block).

Causal relevance

Given that some caution is wanted in counting properties, I suggest that we change our terminology somewhat. Instead of talking about a mental cause M and a physical cause P, let’s talk about events that have various properties – roughly, let’s use the Davidson framework, in which events have many properties. (I do not think that the difference between Davidson-style

events and Kim-style events makes any difference to the mental causation problem.) Instead of saying that a mental event M causes a physical event P*, let us say that the property M is (or is not) *causally relevant* to P*. Causal relevance is thus to be a binary relation on properties. Then, whenever one event *c* causes another event *e*, there are two properties, say X and Y, such that X is causally relevant to Y, and the cause event *c* has X and the effect event *e* has Y.

Intuitively, property X is causally relevant to property Y if one event causes an event that has Y in virtue of or because it has X (Shoemaker, 1984). What shall this mean?

Causal relevance is, in the first instance, to be based on the laws of nature. Mill thought of the individual laws of nature as “threads of connexion” that weave together to produce the “general tissue of connexion” of the natural world, or the fabric of the world. Mackie suggested that we can express the fabric as a collection of necessary and sufficient conditions: a sufficient condition for an instance of a certain property will be a structured array of instances of properties, and the collection of all such sufficient conditions for the effect property is then necessary. Indeterminacy means that some of these conditions will be probabilistic and hence not sufficient, but in such cases there is no sufficient condition to be had. In either case, the conditions will be what Mill called “causes, philosophically speaking” or total causes: a spatial distribution of properties large enough to eliminate all physically possible interferences.

If we focus on a single property of a single occurrence, there is a “swatch” of the fabric of nature that produces that occurrence. So a “swatch law” is a complete specification of what produces an effect of a certain type. The form of such a law is as follows:

For all places *p* and times *t*, if *x* is an event that occurs at place *p* and time *t* and has an array of properties ABCD . . . arranged in a certain structure *S*, then at place *p'* and time *t*+ ϵ there will be an event *y* that has property *G*.

Whenever a “swatch law” of this form is true, then each of the properties mentioned in the antecedent condition is *causally relevant* to the property mentioned in the consequent condition.² To keep the discussion uncluttered, let’s express the “swatch law” this way, leaving the

²Intuition will suggest that a lot of a “cause, philosophically speaking” is causally irrelevant to the effect. For the purposes of this paper I’ll assume that there is a way to articulate what Lewis called our “principles of invidious discrimination” (Lewis, 1973, 162).

structure, time and place in the background:

if x has ABCD . . . , then y has G

The mental causation problem is a problem about the causal relevance of supervening properties. So far, physical properties are causally relevant to other physical properties; what about properties that supervene on the physical properties?

Supervening properties are not “something more”, something “above and beyond”, the physical properties. They are not distinct from the physical properties; nor are they identical to them. They are, rather, supersets of sets of physical properties. So there is room for saying that supervening properties *as well as* physical properties are causally relevant to other properties.

The Millian, laws-of-nature, conception of causal relevance finds causal relevance where there is a necessary and sufficient condition. Lewis’s updated conception of laws of nature would find causal relevance expressed by the predicates of a “best system” of nature. The problem, then, for the causal relevance of supervening properties is that they don’t appear to be connected by laws of nature. There are two obvious reasons why this should seem to be so:

1. The set of fundamental laws of nature is closed in the sense of Kim’s *Closure*. But levels other than the physical aren’t closed; in fact it seems virtually definitive of being a non-physical level that sufficient causes of events at that level involve properties from lower levels.
2. Instances of higher level properties are typically physically different in various ways. The instances of the property of being a cup differ widely: some are porcelain, some are glass, some are wood; and some cups are made of ice. These different kinds of cups differ correspondingly in what they can do. One way to make this point is to say: typically, higher level properties are multiply realizable (but see (Bechtel and Mundale, 1999; Shapiro, 2000) and (Gillett, 2003; Aizawa, 2006)).

The causal relevance of a cup of Earl Grey tea

Let's take a toy example to show how to connect higher level properties by laws of nature: the property of being a cup and its effects.

Cups that are made of porcelain are good for holding hot tea. Those made of ice are no good at holding hot tea but are really good at making messy puddles. For the property of being a cup to be causally relevant to these kinds of effects, that property has to be qualified in a certain way, so that having that property, qualified one way, is causally relevant to holding hot tea, while having that same property, qualified another way, is causally relevant to making messy puddles. What should that qualification look like?

Suppose that at the level of the physical properties, i.e., the first level of causally relevant properties, there are several laws of nature, each making up a "thread" in Mill's fabric of nature, and that they combine into a "swatch" law of that fabric that looks like this:

if x has ABCD . . . , then y has G

Let us suppose that anything that is an ABCD . . . event is an event of pouring hot Earl Grey tea into a porcelain tea-cup. A, B, C, D are physical properties, and each of them is causally relevant to G, another physical property. And let's suppose that T is the property of being a cup. Suppose that T supervenes on property A, so that anything that has A is a cup—as it happens, a porcelain cup. There are other ways to be a cup, so T supervenes on a set that properly includes A. Suppose that any event that has BCD will be an event of pouring hot Earl Grey tea into the cup. Instances of G, then, are instances of holding hot tea.

Now consider the "swatch law"

if x has TBCD . . . , then y has G

that is, the earlier "swatch law", now with 'T' put in for 'A'. Since the instances of T differ physically, this statement is false; sometimes TBCD . . . *isn't* followed by a G. To make it true we will have to ensure that we have the right kind of cup of tea. We do that by adding more properties to the structured array that includes T. We could qualify T with the original A, giving us this condition, which is true if and only if the original condition is true:

if x has TABCD . . . , then y has G

But this qualification is redundant: since being an A is a way to be a T, there is no difference between there being an instance of A, and an instance of TA at the same place and time. (Why couldn't there be *two* ways of being T at the same place and time—the A way, plus some other way of being T? For the same reason that there can't be a completely porcelain cup that is also completely made of ice: the ways of being T are incompatible.) It's also, intuitively, a cheat: this regularity is just the physical regularity, with an irrelevant extra bit added.³

So we need some property other than A to put together with T (and the other elements of the structured array), so that when that collection of properties is instantiated in that structure, the result is an instance of G, and so that the collection isn't redundant.

Something similar to Aristotle's distinction between form and matter will help us out. This particular cup is some porcelain in a certain shape. There are two properties here, one for the form, one for the matter. The form property—T—is the property of being a cup, that is, the property corresponding to the set of all possible cups, made of porcelain, wood, glass, cardboard, titanium, ice, etc. The matter property for this particular cup—call it 'H'—is the property of being *of this kind of stuff*, that is, the property corresponding to the set of all actual and possible arrangements of the exact amount and kind of porcelain of which this cup is made. These arrangements will include ones that are cup-shaped, cube shaped, thin films, etc.

H is a physical property, in the broad sense that all its instances are physical. It is not a property that will be mentioned by the laws of physics, though, so we are not falling back onto the physical laws and the properties they describe.

If T and H are instantiated together in the same thing at the same time, that thing will also instantiate the conjunction of those two properties. Something that has the conjunction of

³(Mackie, 1974, 62) considers the same sort of problem when describing causes as a special sort of sufficient condition:

ABC is a *minimal* sufficient condition: none of its conjuncts is redundant: no part of it, such as *AB*, is itself sufficient for *P*.

Since in our case T supervenes on A, the structured collection TABCD . . . isn't redundant in quite the same way, since by supervenience it's not possible to produce G with that collection *without* T (since it is not possible to produce any instance of the collection ABCD . . . that is not also an instance of T). Nonetheless whatever an A-type T can do is exactly what A-type things can do, so the objection stands.

those two properties is something with the form of a cup, that is moreover made of porcelain: that is, the conjunction of T and H is exactly the original property A. (For this to work, H must be *porcelain* in any of the shapes it can take. It cannot be just any arrangement of the ultimate constituents of a porcelain cup, since those could be rearranged to yield something that is not porcelain.) Then the regularity

if x has THBCD . . . , then y has G

is true if and only if the original,

if x has ABCD . . . , then y has G

is true.

A similar maneuver can be performed on G. Some higher level property, like “holds tea”, supervenes on this physical property. It, like T, can have instances that differ physically. But there are “swatch laws” that generate each of the variations, once the variation is made explicit in the way sketched above.

Finally, the collection of all these laws expresses the laws of nature at the level of the higher level properties. For all the non-higher-level parts of the world, the laws are just the physical laws. For the parts of the world that do have higher level properties, the laws relate suitably qualified higher level properties.

A real example

The question of the causal relevance of higher-level, supervening, properties is the question of whether the properties of the “special sciences” (Fodor, 1974, 1975) are causally relevant. If the properties of a given special science, for instance biology, are not reducible to physical properties, then Kim’s argument would show that they are all causally irrelevant.

A recent discussion in the philosophy of biology uses several of the moves that we have just discussed. Beatty (1995) offers two reasons why biology contains no laws. One is that the fact that there are any biological organisms at all, and that there are biological laws at all, is

contingent on the special sort of initial conditions that obtained (and obtain) on Earth. Change those background conditions, and the laws cease to work. Second, the biological “laws” have exceptions. Mendel’s first “law” says that “diploid sexual organisms form haploid gametes by a “fair” 50/50 meiotic division” (Sober, 1997, S460). But not always, since there are genes that cause segregation distortion. Moreover, Rosenberg (1994) argues that the properties of biology (like fitness) are multiply realized, and hence there can’t be laws governing such properties.

Sober (1997) responds to Beatty’s first reason by pointing out that even if some putative law of biology is contingent on initial conditions, we can make the law non-contingent by including the initial conditions. We can make the same response to Beatty’s second objection: even if some diploid sexual organisms don’t form haploid gametes by a 50/50 meiotic division, there is a difference between those that do and those that don’t. To get exceptionless laws, we just need to include the full characterization of the difference. Finally, the same goes for the worry about multiple realization: to get an exceptionless law, the higher-level property needs to be qualified by additional properties. If we simply added a physical description of the full characterization, we would be characterizing the causes redundantly, as above. But there do exist appropriate “matter” properties which dovetail with the biological properties to yield exceptionless laws.

Properties again

This whole exercise appears to depend on a certain conception of properties and a certain conception of laws and causation. It appears to be pretty much unvarnished Humeanism about cause, law and properties.

But the Humean framework is *not* what is doing the work in this proposal. What is doing the work is the idea that properties can be non-identical while at the same time not distinct. This means that the causal relevance of properties can be *non-additive*.

According to this conception of causal relevance, here is what is going on. Suppose, as Kim would put it, that both M and P cause P*. There are two events, *c* and *e*. *c* has both M and

P. e has P^* . c causes e . It has two causally relevant properties, M and P . This doesn't mean that P^* has two causes, and certainly doesn't mean that there are two independent causes of P^* . It doesn't even mean that c has two distinct properties, each of which is causally relevant to P^* . Counting causes and properties turns out to be less simple than we might have thought; M added to P makes exactly one event, that has two properties that are not completely distinct, yet which also are not identical. (Since distinct = not identical *and* not overlapping.)

The structure of this solution is available for non-Humean accounts of properties, laws and causation. Suppose for example that particular things have particular *dispositions*. A disposition necessitates an effect when joined with reciprocal disposition partners (so Hume was wrong, and there is a "necessary connexion" between being a struck fragile glass, and breaking). The laws of nature are made true by the dispositions that things have.

Now, suppose we say that my wanting the cheesecake is something distinct from the physical basis of the desire. Then there are two distinct dispositions, each of which is sufficient for and necessitates my forming the intention to reach for the cheesecake. But that seems like too many causes, especially given *Closure*, which would seem to entail that the only disposition doing any work is the physical disposition.

But suppose the dispositions are neither identical nor distinct; suppose the dispositions *overlap*. Then there are not two things here. It's not exactly true either that there is one thing or that there is more than one thing.

The non-Humean would now need to say more about *overlap*—does she want to say that one disposition partially constitutes another, or that it is a part of another, or what? She also needs to work out how overlapping dispositions necessitate effects. It is easy enough to say how this works on the Lewisian conception of properties, not so easy on non-Humean accounts. But it's a task worth taking on, since it saves non-reductive physicalism—it yields a solution to the problem of mental causation.

References

- Aizawa, K. (2006). The biochemistry of memory consolidation: A model system for the philosophy of mind. *Synthese*, 155(1):1–34.
- Beatty, J. (1995). The evolutionary contingency thesis. In Wolters, G. and Lennox, J., editors, *Concepts, Theories, and Rationality in the Biological Sciences: The Second Pittsburgh-Konstanz Colloquium in the Philosophy of Science*, pages 45–81. University of Pittsburgh Press, Pittsburgh.
- Bechtel, W. and Mundale, J. (1999). Multiple realizability revisited: Linking cognitive and neural states. *Philosophy of Science*, 66:175–207.
- Fodor, J. (1974). Special sciences (or the disunity of science as a working hypothesis). *Synthese*, 28:97–115. Reprinted in Block (ed.), *Readings in Philosophy of Psychology*, vol. I, pp. 120–133.
- Fodor, J. (1975). *The Language of Thought*. Thomas Crowell, New York.
- Gillett, C. (2003). The metaphysics of realization, multiple realizability and the special sciences. *The Journal of Philosophy*, 100(11):591–603.
- Jubien, M. (1989). On properties and property theory. In Gennaro Chierchia, Barbara H. Partee, and Raymond Turner, editors, *Properties, Types and Meaning: Volume 1*, pages 159–175. Kluwer, Dordrecht.
- Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton University Press, Princeton.
- Lewis, D. (1973). Causation. In *Philosophical Papers*, volume II, pages 159–172. Oxford University Press, Oxford.
- Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, 61(4):343–377.
- Lewis, D. (1986). *On the Plurality of Worlds*. Basil Blackwell, Oxford.
- Mackie, J. L. (1974). *The Cement of the Universe: A Study of Causation*. Clarendon Library of Logic and Philosophy. Oxford University Press, Oxford, paperback edition.
- Rosenberg, A. (1994). *Instrumental Biology or the Disunity of Science*. University of Chicago Press, Chicago.
- Sanford, D. (2005). Distinctness and non-identity. *Analysis*, 65(4):269–74.
- Shapiro, L. A. (2000). Multiple realizations. *The Journal of Philosophy*, 97(12):635–54.
- Shoemaker, S. (1984). Causality and properties. In *Identity, Cause and Mind: Philosophical Essays*, pages 206–233. Cambridge University Press, Cambridge.

Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, 68(2):141–56.

Sober, E. (1997). Two outbreaks of lawlessness in recent philosophy of biology. *Philosophy of Science*, 64 Supplement:S458–S467.

Taylor, B. (1993). On natural properties in metaphysics. *Mind*, 102(405 (January)):81–100.