Karen Bennett on causal exclusion: A nonreductive physicalist solution

Introduction

Recall from last time that our five claims – Completeness, Efficacy, No Systematic Overdetermination, and Exclusion, and Distinctness – make an inconsistent set. How to respond?

Today we will consider a nonreductive physicalist solution that Karen Bennett has developed in a series of articles. Along the way, we will also consider some specific versions of nonreductive physicalism in more detail than we have to date.

Introduction

Like many contemporary physicalists, Bennett is a nonreductive physicalist. She takes this to involve acceptance of the first four claims:

> Completeness, because she thinks the universe is a closed system of physical causation.
>
> Efficacy, because she wants to maintain our common-sense belief that thinking and feeling can make things happen in and to our bodies.
>
> No Systematic Overdetermination, because she thinks it is wildly implausible to hold that any time the mental interacts with the physical there is causal overdetermination (of the firing squad variety).
>
> Distinctness, because multiple realizability considerations lead her to reject reductive physicalism (even though she does think that the mental supervenes on the physical).

She thinks she can have all of this (and eat her cake too) by denying the fifth claim:

> Exclusion—No effect has more than one sufficient cause unless it is causally overdetermined.

Introduction

Wait. What? How could an effect have two distinct sufficient causes but fail to be overdetermined?

Introduction

Bennett's answer: if the two sufficient causes are tightly metaphysically connected in the right sort of way, then, while they are each sufficient causes of the same effect, they don't overdetermine that effect. By making the mental metaphysically supervene on the physical, nonreductive physicalists posit just such a tight metaphysical connection between mental events and physical events. So nonreductive physicalists (at least) can allow joint sufficient causes without overdetermination: Exclusion is false!

- Another thing you want to make sure you understand: As a nonreductive physicalist, Bennett believes that any possible world that is a minimal physical duplicate of the actual world is a duplicate simpliciter. This entails that the actual way things are physically metaphysically necessitates the actual way they are mentally. (This will be important for understanding why Bennett thinks the property dualist is not in a position to adopt the same solution to the causal exclusion problem.)

- If Bennett is right, then the causal exclusion *is* a genuine problem, but only for dualists.

Overdetermination: a necessary condition

Bennett rests her argument on a very plausible necessary condition on overdetermination. Before getting into the weeds, let's illustrate the key point intuitively.

Billy and Suzy and are throwing rocks at windows. Suppose that in one case the window shatters and the shattering of the window is causally overdetermined by Billy's throw and Susy's throw, such that each of their throws was fully sufficient to shatter the window and that neither of their throws has a special claim over the other to being "the" cause of the window shattering.

Overdetermination: a necessary condition

Intuitively, this seems to imply that had Susy *not* thrown her rock but Billy still had thrown his, the window would *still* have shattered, and vice versa: had Billy *not* thrown his rock but Susie had thrown hers, the window would *still* have shattered.

Bennett's idea: this feature of the Billy/Susy case is a *fully general* feature of causal overdetermination, including of a physical effect's being (putatively) causally overdetermined by a mental event and by physical events. The non-vacuous truth of both counterfactuals is (at least) a *necessary condition* for a genuine causal overdetermination.

If so, then an attractive strategy suggests itself for nonreductive physicalists for responding to the causal exclusion problem: argue that if nonreductive physicalism is true, then this necessary condition for causal overdetermination isn't satisfied in cases of mental causation.

Now, into the weeds …

Overdetermination: a necessary condition

Let m be a mental cause at t_1, p be a physical cause at t_1, and e be a physical effect at some later time t_2. And suppose that m and p are each sufficient causes of e. We can now express our intuitively necessary condition on overdetermination in terms of the (non-vacuous) truth of two counterfactual conditionals about m, p, and e:

(O1) If m had happened without p, e would still have happened: $(m \,\&\, \neg p) \,\Box\!\!\rightarrow e$.

(O2) If p had happened without m, e would still have happened: $(p \,\&\, \neg m) \,\Box\!\!\rightarrow e$.

If e is genuinely causally overdetermined by m and p, then both (O1) and (O2) are non-vacuously true.

Bennett's task: demonstrate that if nonreductive physicalism is true, then this condition doesn't obtain – i.e., at least one of (O1) or (O2) is not non-vacuously true. There are two ways that a conditional can fail to be non-vacuously true: it can be false or it can be vacuously true. So, Bennett's task is to demonstrate that if nonreductive physicalism is true, then either (O1) or (O2) is false or vacuous.

Logical interlude about counterfactual conditionals (this will be a little technical)

(O1) If m had happened without p, e would still have happened: $(m \,\&\, \neg p) \,\Box\!\!\rightarrow e$.

(O2) If p had happened without m, e would still have happened: $(p \,\&\, \neg m) \,\Box\!\!\rightarrow e$.

The standard semantics for counterfactual conditionals makes them vacuously true *whenever they have a necessarily false (i.e., impossible) antecedent*. To see why this is, we need to grasp the standard semantic analysis of counterfactual conditionals:

$P \,\Box\!\!\rightarrow Q$ iff $\forall w(SIM(P(w)) \rightarrow Q(w))$

["If P had been the case, then Q would have been the case" iff for all possible worlds most similar to the actual world at which P is true, Q is also true"]

- This standard analysis employs quantification over possible worlds in order to treat counterfactuals as having the logical form of a classical material conditional ('$\rightarrow$').

Logical interlude about counterfactual conditionals (this will be a little technical)

Let's review the truth table for the material conditional (one of the classical truth-functional logical connectives, together with '~' ('not'), '&' ('and'), 'V' ('disjunction'), and '↔' ('biconditional').

A material conditional (e.g., P → Q) with a false antecedent (P) is vacuously true (i.e., true whether Q is true or false.)

P	~ P
T	F
F	T

P	Q	P & Q
T	T	T
T	F	F
F	T	F
F	F	F

P	Q	P V Q
T	T	T
T	F	T
F	T	T
F	F	F

P	Q	P → Q
T	T	T
T	F	F
F	T	T
F	F	T

Logical interlude about counterfactual conditionals (this will be a little technical)

A lot of people feel like they intuitively grasp the truth tables for the other truth-functional logical connectives: '~' ('not'), '&' ('and'), and 'V' ('disjunction'), but struggle to grasp the truth table for '→' ('if…then…'), at least at first.

Imagine that you are a security guard whose job is to check whether all patrons of the establishment at which you work are following the local liquor law. We can express the liquor law as a material conditional: that *if* X is drinking alcohol, *then* X is 19 years or older. To establish whether the liquor law is being obeyed (i.e., whether it is *true* that everyone who is drinking alcohol is 19 years or older), who would you need to check? Let's run through the possibilities …

X is drinking alcohol	X is 19 years or older	X is drinking alcohol → X is 19 years or older
T (e.g., X is drinking beer)	T (e.g., X is 45 years old)	T
T (e.g., X is drinking vodka shots)	F (e.g., X is 10 years old)	F
F (e.g., X is drinking water)	T (e.g., X is 75 years old)	T
F (e.g., X is drinking water)	F (e.g., X is 12 years old)	T

P	Q	P → Q
T	T	T
T	F	F
F	T	T
F	F	T

Logical interlude about counterfactual conditionals (this will be a little technical)

Now, let's return to the standard semantics for counterfactual conditionals:

$$P \mathrel{\square\!\!\rightarrow} Q \text{ iff } \forall w(SIM(P(w)) \rightarrow Q(w))$$

["If P had been the case, then Q would have been the case" iff for all possible worlds most similar to the actual world at which P is true, Q is also true"]

- Since the standard semantics of the counterfactual conditional treats the counterfactual conditional as a special case of the material conditional, we can see why a counterfactual conditional with *a necessarily false (i.e., impossible) antecedent is vacuously true*. Suppose that P is necessarily false (i.e., that there is no possible world at which P is true). If so, then it follows that no possible world that is similar to the actual world is a world at which P is true, and this means that the antecedent of the material conditional on the right-hand side of our above analysis is *false*. But, as we've seen, according to the truth conditions for a material conditional, whenever the antecedent of a material conditional is false, the material conditional itself is vacuously true. Given that, and given that we have analyzed counterfactual conditionals in terms of material conditional, we obtain the result that if P is necessarily false, then any counterfactual conditional of which P is the antecedent is merely vacuously true.

Karen Bennett on causal exclusion: A nonreductive physicalist solution

Overdetermination: a necessary condition

The question confronting nonreductive physicalists (according to Bennett): on their view, is either of the following counterfactual conditionals false or vacuously true?

(O1) If m had happened without p, e would still have happened: $(m \,\&\, \neg p) \mathrel{\Box\!\!\rightarrow} e$.

(O2) If p had happened without m, e would still have happened: $(p \,\&\, \neg m) \mathrel{\Box\!\!\rightarrow} e$.

The status of (O1)

Bennett says that "the status of O1 is complicated for the physicalist" (380). She expands on why in the optional reading for today (see Bennett 2003: pp. 481-484). There, Bennett argues that it is dialectically stronger for nonreductive physicalists to focus on O2 because it speaks more directly to the worry that nonreductive physicalists will end up having to deny Efficacy, given their commitment to Distinctness, Completeness, and No Systematic Overdetermination (p. 2003: 490).

So, let's follow Bennett in focusing on O2 and ignoring O1. If she can show that O2 is either false or vacuously true, we won't need to worry about the status O1.

The status of (O2)

What can nonreductive physicalists say about O2? According to Bennett, what they say will depend on what they think it is for an event like p to be "causally sufficient" for an effect like e.

- If what is causally sufficient for e is p, strictly defined, *together with a whole host of complicated extrinsic factors regarding all the background conditions for p's occurren*ce, then O2 is vacuous. This is because, given supervenience physicalism, p + background conditions *metaphysically necessitates* m. Accordingly, the antecedent of O2 (p & ¬m) is impossible. And a counterfactual conditional with an impossible (i.e., necessarily false) antecedent is vacuously true.

- If, on the other hand, what is causally sufficient for e is *simply* p, strictly defined, then O2 is false. This because no one thinks (or anyway should think) that p, in utter isolation from all background conditions (e.g., some brain cells a petri dish) is sufficient for much of anything, let alone e.

The status of (O2)

The upshot (according to Bennett): whichever way you define "causal sufficiency", nonreductive physicalists have a plausible way to claim that O2 is either vacuously true or false. If so, then Exclusion is false: we have a case in which physical effect, e, possesses two sufficient causes, m and p, and yet the case at hand doesn't satisfy our requirement on overdetermination.

In intuitive terms, mental causation, for the nonreductive physicalist, is nothing like death by a firing squad.

Objection

Has Bennett not lapsed into reductive physicalism? Specifically, we might grant her that p (strictly defined) is distinct from m. But what reason is there to not reductively identify m with p + background conditions – i.e., with the physical event that *does* metaphysically necessitate m?

- Bennett doesn't address this question in as much detail as she might have (possibly because she wants to remain neutral between different versions of nonreductive physicalism). To see how she and other non-reductive physicalists could reply to this challenge, it helps to consider a few specific nonreductive physicalist accounts of the relation in virtue of which mental properties metaphysically supervene on physical properties.

The determinable-determinate strategy (Yablo)

Suppose Sophie the pigeon is trained to peck only when she is presented with a decent-sized red patch. You present her with a scarlet patch, and (predictably) she pecks. What caused her pecking? On the one hand, it seems to be the instantiation of redness (a 'determinable' property) that causally explains Sophie's pecking. On the other hand, in pecking at a red patch, Sophie is also pecking at a red patch of a 'determinate' shade of red (namely, scarlet). So, scarlet is also a sufficient cause of her pecking.

But notice: these two causal explanations of Sophie's pecking are perfectly compatible. The determinable property of being red and the determinate property of being scarlet are not "causal rivals" with respect to Sophie's pecking, and so her pecking isn't overdetermined. And yet the patch's being red and its being scarlet are each causally sufficient for Sophie's pecking. Upshot: Exclusion is false wherever we find causation by a determinable property!

- On Stephen Yablo's version of nonreductive physicalism, every mental property, M, is a determinable of some determinate physical property, P_1, P_2, … P_n.

- Stephen Yablo "Mental Causation" (1992)

The causal subset strategy (Wilson, Shoemaker)

According to Wilson (optional reading), underlying Yablo's determinable-determinate strategy (as well as all other versions of nonreductive physicalism) is a more fundamental truth: that the causal powers of a higher-level (e.g., mental) property are a *proper subset* of the causal powers of the lower-level (e.g., physical) property that realizes the higher-level property. (We can illustrate this with Sophie the pigeon).

What makes Distinctness true, on Wilson's version of nonreductive physicalism, is that mental properties have *fewer* causal powers than their realizing physical properties. The relation in virtue of which the mental metaphysically supervenes on the physical is the relation of proper parthood. Mental properties *just are* a causal subset of physical properties.

- See Jessica Wilson, "How Superduper does a Physicalist Supervenience Need to be?" (see also her *Metaphysical Emergence*, 2019)

Answering the objection?

Exercise: Using one of these nonreductive physicalist proposals, show how Bennett could answer the challenge we raised earlier against her proposed solution to the causal exclusion problem.

A possible Yablo-inspired answer: p + background conditions ≠ m because p + background conditions is a determinate of the determinable m.

A possible Wilson-inspired answer: p + background conditions ≠ m because the causal powers of m are a proper subset of the causal powers of p + background conditions.

How do things stand for the dualist?

Can dualists do the same thing that Bennett has said on behalf of the nonreductive physicalist? That is, can they keep Distinctness, Completeness, Efficacy, and No Systematic Overdetermination while escaping contradiction by rejecting Exclusion?

Bennett argues that they can't. The causal exclusion problem seems to be a *genuine* problem for dualists.

- Importantly, Bennett *restricts* her remarks to dualists who want to avoid epiphenomenalism, violations of the completeness of physics, and systematic overdetermination. So if you are a dualist who is prepared to embrace one of these claims, she isn't talking to you.

How do things stand for the dualist?

On O1: Dualists must say that O1 is non-vacuous since dualism entails precisely that m can happen without p. But the dualist must also say that O1 is true. Why? Because saying that m can't bring about e without p contradicts the assumption that m is causally sufficient for e.

On (O2): Dualists must say that O2 is non-vacuous since dualism entails precisely that p can happen without m. But the dualist must also say that O2 is true. Why? Because saying that p can't bring about e without m contradicts the assumption that p is causally sufficient for e.

How do things stand for the dualist?

Upshot (according to Bennett): Dualists can't deal with the problem the way the nonreductive physicalist can: namely, by denying Exclusion. They are stuck with having to deny Completeness, Efficacy, or No Systematic Overdetermination.

Questions to think about:

- Is Bennett right about this?
- If she is right, how much of a problem is this for the dualist? Why shouldn't dualists reject, say, Completeness?