

The Overdetermination Argument versus the Cause-and-Essence Principle—No Contest

PAUL NOORDHOF

Scott Sturgeon has claimed to undermine the principal argument for Physicalism, in his words, the view that “actuality is exhausted by physical reality” (Sturgeon 1998, p. 410). In noting that actuality is exhausted by physical reality, the Physicalist is not claiming that all that there is in actuality are those things identified by physics. Rather the thought is that actuality is made up of all the things identified by physics and anything which is a compound of these things. So there are tables as well as their micro-physical constituents. The argument that Sturgeon has in his sights is the Overdetermination Argument. In what follows, I shall argue that Sturgeon’s criticism of the Overdetermination argument fails. I shall also argue that Physicalism can accommodate his claim that causal statements concerning the mental and physical respectively may require diverse patterns of counterfactual activity for their truth.

1. Overdetermination Argument

The Overdetermination Argument for Physicalism runs as follows:

- (1) Every physical effect has its chance fully determined by physical events alone.¹
- (2) Mental events have physical effects.
- (3) The physical effects of mental events are not generally overdetermined.

Therefore,

- (4) Mental events are physical events (see Sturgeon 1998, pp. 413–4).

Sturgeon suggests there is an equivocation that vitiates the argument as it stands. In premise (1), the term physical means that which is identified by physics. So, if Quantum Mechanics is correct, it amounts to the claim that

¹ I have used this formulation drawn from Sturgeon’s elucidation of his preferred characterisation—namely every physical effect has a fully revealing, purely physical history—because it seems more perspicuous (see Sturgeon 1998, p. 413).

Every quantum effect has its chances fully determined by quantum events alone.

Understood this way, the claim has some degree of plausibility. However, it is no part of physics or any other science that mental events have physical effects in this sense. So the sense of physical in premise (2) is broader. It should refer to effects identified by special sciences like geology as well as everyday physical effects like the knife being knocked off the kitchen table. Understood this way, Sturgeon questions whether any science claims that every physical effect has its chances fully determined by physical events alone. He suggests that this cannot be the case because both some special sciences and everyday experience rely upon *mental* causes for broadly physical effects. The question is whether the equivocation can be banished and the two premises be true under one sense of the physical (see Sturgeon 1998, pp. 415–6).

2. *Transmission of causality*

The connection between the broad notion of the physical and that defined in terms of physics is that

Broadly physical events are composed of quantum physical events. (Sturgeon 1998, p. 417)

Is there a way to appeal to this fact in order to explain how the efficacy of the mental and the physical overdetermine behaviour? Initially it might seem so. Sturgeon considers the following composition principles which indicate how this might be done.

Downward:

If C causes E and E is composed of $E_1, E_2 \dots E_n$, then C causes each E_i ;

Upward:

If C causes some E_i and $E_1, E_2 \dots E_n$ compose E , then C causes, E [where $1 \leq i \leq n$ and E_i is an event which partly constitutes E].²

He finds them wanting. He provides a number of counterexamples of what he considers to be increasing strength. The weakest of these, by his reckoning, is perhaps the clearest to make the point that these principles are inadequate. Since I propose to concede this point for the sake of argument, I will not go through the others (see Sturgeon 1998, p. 419).

² I have reformulated the principles found in Sturgeon to make explicit his elucidation of the principles.

1000 ducks are on a lake. All are normal save Duck₁₀. Duck₁₀ is deaf. As it happens, Duck₁₀ is bitten by a turtle just as a shotgun is fired nearby. The flock takes flight en masse. (Sturgeon 1998, p. 419)

Intuitively, the turtle does not cause the flock to fly even though it causes one of the flock to fly. Hence there is a counterexample to Upward transmission. Intuitively, the shotgun blast does not cause Duck₁₀ to fly even though it does cause the flock to fly. Hence there is a counterexample to Downward transmission.

The natural diagnosis of what is wrong is that the flight of Duck₁₀ is not essential to the flock flying. So causing the flock to fly does not *require* causing the flight of Duck₁₀. Likewise, as the flight of Duck₁₀ is not a necessary condition of the flock flying, causing the flight of Duck₁₀ does not cause the flock to fly. Examples like this motivate the *Cause-and-Essence Principle*:

C causes *E* iff *C* is sufficient to bring about what's essential to *E*. (Sturgeon 1998, p. 422)³

Sturgeon appeals to this principle to block the Overdetermination Argument, or so it first appears. However, on reflection, it seems that he is appealing to something else which seems rather less well motivated and which clashes with his formulation of the Cause-and-Essence Principle. Once these matters are extricated, it becomes apparent that he has not blocked the Overdetermination Argument.

Sturgeon notes that there is little conceptual gap between our notion of the flight of a flock of ducks and the flights of individual ducks. He suggests that it is the absence of the conceptual gap that makes this counterexample to the composition principles weak. By contrast, he considers there is a significant conceptual gap between the decay of an atom and selling on the stock market even if the first partly constitutes the second (Sturgeon 1998, pp. 421–3). It is this which makes a counterexample with a similar structure involving the decay of an atom, selling stock and the stock market crash seem stronger. Of course, there is doubtless a way in which the first counterexample is weaker and the second stronger (perhaps in dramatic appeal or comparative size of the macro–micro constituents). However, I don't understand how this relates to Sturgeon's diagnosis of the problem with the composition principles. The Cause-and-Essence Principle shows why the Duck case is just as much a counterexample as the subatomic case. Even if there is little conceptual gap in the Duck case, there

³ Two qualifications seem necessary to this principle. First, *C* should just be sufficient to fix the probability of what's essential to *E* occurring (to account for indeterministic causation). Second, *C* needn't be sufficient to bring about all of what's essential to *E*, only part. However, these qualifications are not necessary for the argument that follows.

is still no transmission because of the Cause-and-Essence Principle. Conversely, I think that Sturgeon is right to insist that the Cause-and-Essence Principle provides a sufficient condition for the transmission of causality even if there is a conceptual gap. He does not suggest that if there is a large conceptual gap between say micro-event C and macro-event E , then C is not a cause of E . He suggests that C would be a cause of E if it were sufficient to bring about what is essential to E .

The problem is that when Sturgeon seeks to explain why mental events and microphysical events don't overdetermine behaviour he appeals to the conceptual gap rather than the Cause-and-Essence Principle. His case rests upon two disturbing features of Quantum Mechanics. First, there is *superposition*. When their position is not measured, particles are best conceived of as wave-like phenomena which have a certain probability of being at various positions. Thus, suppose that it is equally probable that the particle is located at P_1 or P_2 or P_3 , then it can be characterized as a probability wave $1/3$ at P_1 , $1/3$ at P_2 , $1/3$ at P_3 . By contrast, our conception of everyday macro-objects is that they don't have this feature. If they have an equal probability of being at three places, that does not mean that they are $1/3$ at each. Second, there is *projection*. When the position of the particle is measured, the wave, which is one third at each of the positions indicated, collapses into a particle in a particular position. Once more, our conception of macro-objects is that this is not the case for them. I am prepared to concede that quantum mechanics ascribes to microphysical objects the features Sturgeon identifies (in fact, I am in no position to debate this matter). Also, regardless of whether Sturgeon is exactly right about this, I think he is right that there is this gap between how quantum mechanics characterizes things and how our everyday conception of macro-objects says that they are. However, I do not see how pointing to this gap shows that causes of macro-objects aren't causes of quantum effects or causes of quantum effects aren't causes of macro-objects. What justified our intuition that it was inappropriate to transmit causality up or down was not the gap in our conception but the Cause-and-Essence Principle. So it seems to me that the crucial issue is whether the Cause-and-Essence Principle proclaims that causality can't be transmitted in this case.

Looked at from this perspective, there seems to be no problem with the transmission both up and down so long as they satisfy one of the following principles.

Downward:

If C causes E and E is *essentially* composed from $E_1, E_2 \dots E_n$, then C causes at least one of $E_1, E_2 \dots E_n$;

Upward:

If C causes at least one of $E_1, E_2 \dots E_n$ and $E_1, E_2 \dots E_n$ essentially compose E , then C causes E .

These incorporate the insight captured by the Cause-and-Essence principle. Take my arm movement. Sturgeon has accepted that my arm movement is composed of quantum physical events (Sturgeon 1998, pp. 417, 428). All we must do is identify the minimum number of quantum events necessary for my arm to move and consider whether these are overdetermined by both quantum events and a mental event. Surely the answer is yes. If my arm movement is genuinely composed of quantum events then, if a mental event did not cause the minimum number of quantum events, it could not move my arm. Likewise, since quantum events serve to cause each of the quantum events which make up the minimum required for the arm movement to occur, it follows that collectively they cause the arm movement. So we seem to have found overdetermination after all. Sturgeon writes

Since it's unclear how macro movements spring from quantum events, it's unclear that causes of the latter thereby cause the former. (Sturgeon 1998, p. 427)

That's fine, so long as one does not agree that macro-movements are composed from quantum events. If Sturgeon's argument were that we have no reason to believe that macro-movements are composed from quantum events because of the conceptual gap, I could understand his position. He claims that they are so composed:

it's a puzzling fact that quantum events compose into duck-movements [and presumably arm movements]. We know they do somehow, but we do not know how. (Sturgeon 1998, p. 427)

What I find puzzling is the combination of this claim with his resistance to what seems to be a consequence. If one thinks that macro-movements are so composed, then although it is unclear *how* macro-movements spring from quantum events (or vice versa), it is clear *that* they do.⁴

My conclusion rests on the assumption that a certain minimum number of quantum events is essential to the presence of an arm movement. Is this assumption correct? I could stick with an ad hominem point. Sturgeon commits himself to component essentialism for aggregates—the idea that

⁴ Sturgeon also suggests that interpretations of quantum mechanics indicate that other premises of the Overdetermination argument are less well supported than people think (Sturgeon 1998, pp. 427–8). I don't propose to discuss these in any detail except to note that

- (a) I agree that Premise (1) is false if Wigner is right that the wave function collapses just because of the activity of irreducibly non-physical mental properties.

the components of aggregates are essential to aggregates (Sturgeon 1998, p. 425, fn.13). He also seems to commit himself to arm movements being aggregates. However, it would be nice to do better in response to this objection than point to Sturgeon's commitments. So here are two replies. First, what I am saying is essential to the presence of an arm movement is not all its components but just the presence of a minimum number of its component quantum events. This does not mean that any particular quantum event is essential to the presence of the arm movement. I claim that the presence of a minimum number of component quantum events is caused both by the arm movement and by an antecedent ensemble of quantum events. This response preserves the Cause-and-Essence principle but countenances the idea that quantum mechanics can identify events that fix the chance of the event of there being a minimum number of quantum events present.

Suppose that not even the presence of a minimum number of quantum events is essential to a particular arm movement. Perhaps *my* arm movement (say) could occur in a very different possible world where there are no quantum events. However, that doesn't mean that something fails to be sufficient to bring about what's essential to my arm movement—it being an arm movement say—by just being a cause of the minimum required number of quantum events for the arm movement to occur in the actual world. To establish this, one would have to read “bring about” in the Cause-and-Essence Principle as “cause”. In which case, the principle is in need of reformulation. It should read

C causes E iff C is sufficient to bring about what's essential for E to occur in the circumstances. (Sturgeon 1998, p. 422)

In the actual world, my arm movement can't occur without a minimum number of quantum events. Any more would be unnecessary. How my arm movement might occur in other (far distant) possible worlds is irrel-

-
- (b) Premise (1) appears unthreatened by Bohr's interpretation because even if the collapse of the wave packet occurs because of the interaction between quantum and classical systems, it is not part of the justification of premise (1) that quantum mechanics is complete, just that physics is. I agree that Bohr's interpretation tells against the claim that every quantum spatial effect has a fully revealing, purely quantum history but question whether the overdetermination argument should be formulated in terms of this more specific claim.
 - (c) Premise (2) appears unthreatened by the many minds interpretation of Everett because even if it is true that mentality does not cause the collapse of the wave function, it still causes behaviour.

So it does not seem to me that Sturgeon's discussion of the measurement problem adds significantly to the case he has made against the Overdetermination Argument.

evant. If this is right, the composition principles should be modified accordingly.

3. Physicalism, causality and close possible worlds

Sturgeon claims that counterfactual statements like

- (a) If I had not seen her, I would not have waved my arm
- (b) If quantum events E_1 and E_2 had not occurred, I would not have waved my arm

invoke different sets of closest possible worlds in virtue of which they are true (see Lewis 1979).

He seems to suggest that those who use the overdetermination argument to defend Physicalism by appealing to the composition principles canvassed above lose the capacity to recognize this. If Reductive Physicalism is true, then mental properties are properties identified by completed Physics. In this case, Sturgeon looks to be right. Putting it crudely, we might take the property of seeing her to be identical to the property instantiated as the occurrence of E_1 and E_2 . In which case, the same sets of possible worlds appear to be invoked. However, I do not think that Sturgeon is right if Non-Reductive Physicalism is true. Non-Reductive Physicalism arguably allows for the following cases.

- (i) The property of seeing her is variably realised by physical properties (sometimes by the property instantiated as the occurrence of E_1 and E_2 and other times by a range of other physical properties).
- (ii) The property of seeing her may be realized also by non-physical properties. It is just that it is not so realized in our world.

These cases have ramifications for the assessment of the counterfactuals above. I shall focus on (i) which is less contentious.

If we consider circumstances in which E_1 and E_2 had not occurred, then it is not ruled out that some other realization of the property of seeing her may be present. There is no reason to suppose that some worlds in which another realization is present will fail to be among the closest possible worlds in which E_1 and E_2 do not occur. So counterfactual (b) may be false. By contrast, if the property of seeing her is not present, then not only is the occurrence of E_1 and E_2 excluded, but also the occurrence of all other realizations of the property. Therefore, it is plausible that in the closest possible worlds relevant for this case, no causes of my waving will be present. So counterfactual (a) may be true. But even if this is not so, the possible worlds in which a cause of my waving is present will be ones in

which something other than the property of seeing her is realized. This will be a very different set of worlds from the ones we envisaged to be relevant to the assessment of counterfactual (b)—the one concerning the quantum events.⁵

The consequence of this point is that the Non-Reductive Physicalist can give the same answers to Sturgeon's two questions as the Non-Physicalist. Taking them in reverse order, Sturgeon's second question was

Are the patterns (i. e. sets of the closest possible worlds) indexed to macro-causal claims identical to those indexed to micro-causal claims?

I have just illustrated in my discussion of the example above how these sets of possible worlds may be different. Sturgeon's first question was

Are the regions (of modal reality) indexed to macro-causal claims *contained within* those indexed to micro-causal claims?

The question here concerns whether the set which is the conjunction of all the sets of closest possible worlds in which we assess micro-causal claims is distinct from the set which is the conjunction of all the sets of closest possible worlds in which we assess macro-causal claims. As far as I can see, there is no reason to think that these sets will be identical given the type of case described above. If the property of seeing her is not instantiated in another possible world, this may be because certain psychological and psychophysical laws of the actual world no longer hold. It may involve significant differences of physical law too. By contrast, if E_1 and E_2 are not present, the differences of physical law might be rather less. This shows that the sets of closest possible worlds required to assess relatively simple counterfactuals involving relatively slight changes are distinct for macro- and micro-causal statements. There is no reason to think that these differences would even out over the entirety of macro- versus micro-causal claims. So the Non-Reductive Physicalist can endorse the composition principles, allow that both the instance of the property of seeing her and the quantum event pair are causes, and yet capture things Sturgeon wishes to say about the character of causality.

4. Conceptual gaps and concept-free causation

My objection to Sturgeon's argument rests on my failure to perceive the strength of Sturgeon's remarks about how the presence of a conceptual gap affects attributions of causality. Perhaps this is a mistake. Perhaps we

⁵ For more detail concerning how Non-Reductive Physicalists should conceive of efficacy at macro and micro levels see Noordhof (forthcoming).

should modify the Cause-and-Essence Principle so that it only provides a necessary condition and have a clause requiring the absence of a conceptual gap. But I think that if causality can only be attributed in the way Sturgeon implicitly recommends, we should cease to talk of causality but talk of a notion that ignores the presence of a conceptual gap in determining its attributions—call it concept-free causality. According to concept-free causality—which may be what Sturgeon has in mind by the notion of “inducing”—if mental events were distinct from physical events, there would be concept-free overdetermination (Sturgeon 1998, pp. 428–30). This still seems to be a significant problem. The presence of systematic concept-free overdetermination appears to remain an unattractive metaphysical upshot of the Non-Physicalist’s position. Successful opposition to the Overdetermination argument requires a defence of the claim that the overdetermination-surrogate that is still allowed is much better than the overdetermination that was detected before.

Department of Philosophy
University of Nottingham
University Park
Nottingham NG7 2RD
UK
Paul.Noordhof@nottingham.ac.uk

PAUL NOORDHOF

REFERENCES

- Lewis, David 1979: “Counterfactual Dependence and Time’s Arrow”, in his *Philosophical Papers*, Volume 2, Oxford: Oxford University Press, 1986, pp. 32–66. Originally published in 1979 in *Noûs*, 13.
- Noordhof, Paul (forthcoming): “Causation by Content?” *Mind and Language*.
- Sturgeon, Scott 1998: “Physicalism and Overdetermination”. *Mind*, 107, pp. 411–32.

