

Causal overdetermination for Humeans?

Michael Esfeld, University of Lausanne, Department of Philosophy

Michael-Andreas.Esfeld@unil.ch

(forthcoming in *Metaphysica* 11 (2010), issue 2)

Abstract

The paper argues against systematic overdetermination being an acceptable solution to the problem of mental causation within a Humean counterfactual theory of causation. The truth-makers of the counterfactuals in question include laws of nature, and there are laws that support physical to physical counterfactuals, but no laws in the same sense that support mental to physical counterfactuals.

Keywords counterfactuals, Humeanism, laws, mental causation, overdetermination

1. *The overdetermination option*

The problem of the causal efficacy of mental properties – or the properties in which the special sciences trade in general – consists in the fact that the conjunction of the following four principles is inconsistent:

- 1) *Distinction*: Mental events are not identical with physical events.
- 2) *Causation*: Mental events cause physical events.
- 3) *Completeness*: Any physical event has sufficient physical causes, insofar as it has causes at all.
- 4) *No systematic overdetermination*: Events are not systematically overdetermined by mental and physical causes.

Let an event be the instantiation of a property by an object, such that a mental event is the occurrence of a mental property, and a physical event the occurrence of a physical property. Of course, a more precise formulation of these principles is needed in order to prove the inconsistency of their conjunction. However, since no one disputes the fact of there being such an inconsistency, let this rough and ready formulation do for present purposes.

Since Kim (1998) has drawn new attention to the problem of mental causation, the discussion has mainly focused on (1). In recent years, however, a number of authors have claimed that it is reasonable to abandon (4), provided that one endorses a Humean theory of causation (see notably Bennett 2003 and Loewer 2007b as well as Marras 2007, pp. 318-319; Kroedel 2008; Harbecke 2008, chapter 4). In that way, the metaphysics of causation becomes pertinent to the problem of mental causation. It is common to draw a distinction between dependence theories and production theories of causation (Hall 2004). Humean theories of causation, be they simple regularity theories, be they sophisticated regularity theories in terms of counterfactual dependence, are the most prominent example of the former.

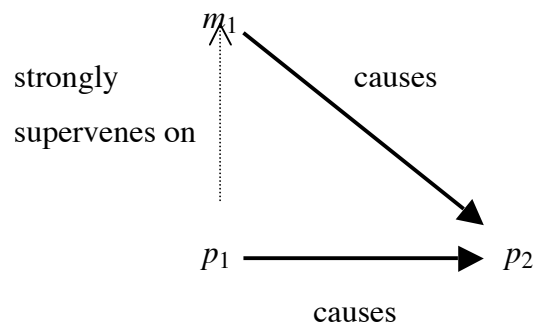
The claim thus is that by settling for a dependence theory of causation in contrast to a production theory, one makes a solution to the problem of mental causation available that consists in rejecting (4) by admitting systematic overdetermination: whenever a mental event – or any event that consists in the instantiation of a property in the domain of a special science – causes a physical event, the effect is overdetermined by both a sufficient mental and a

sufficient physical cause. In the discussion with Loewer (2007b), Kim (2007) commits himself to a version of the production theory of causation, and moreover formulates his view in such a way that it may cohere with classical mechanics, but obviously can work neither for general relativity theory nor for quantum theory, thus provoking the additional objection that the whole problem rests on an outdated conception of physics (Ladyman 2008).

The claim of this short paper is that this turn of the discussion is misguided: whatever the current or future best physical theory may tell us about the world and whatever stance one takes in the metaphysics of causation, one faces the problem of mental causation. In particular, there is no easy way out of this problem for Humeans by simply abandoning (4).

2. *The relevance of laws*

Consider a situation in which a mental event m_1 causes a physical event p_2 by (2), with p_2 having also a sufficient physical cause p_1 by (3), while m_1 is not identical with p_1 (by (1)), so that there is a case of overdetermination, contradicting (4) if such a scenario holds for mental causation in general. Assume furthermore that mental properties strongly supervene on physical properties:



One would have to set out a precise definition of strong supervenience, but we can again skip such complications for present purposes: the only function that strong supervenience has in this context is to ensure that there is no possible world in which there is an event of the same type as p_1 without there also being an event of the same type as m_1 . Consequently, in the context of a counterfactual theory of causation, if p_1 causes p_2 , the proposition “If p_1 had not occurred, p_2 would not have occurred either” is a true counterfactual. Then, by strong supervenience, the proposition “If m_1 had not occurred, p_2 would not have occurred either” is a true counterfactual as well. However, the mere fact of there being a true counterfactual that links the absence of p_1 with the absence of p_2 and there being a true counterfactual that links the absence of m_1 with the absence of p_2 , while strong supervenience ensures that an event of the same type as p_1 always goes together with an event of the same type as m_1 , is not sufficient to establish that there is a causal relation between m_1 and p_2 , not even on Humean standards.

According to Humean metaphysics, causal relations supervene on the distribution of the fundamental physical properties throughout the whole of space-time, these being perfectly natural and categorical properties (see Lewis’ thesis of Humean supervenience in e.g. Lewis 1986a, introduction). Nonetheless, the truthvalue of counterfactual propositions about the actual world is fixed through similarity relations between the actual world and other possible worlds (see notably Lewis 1973 and 2004), without that manner of fixing their truthvalue

necessarily committing the Humean to realism about other possible worlds (see Loewer 2007a). Again, the details of how the truthvalue of counterfactual propositions about the actual world is fixed through similarity relations with other possible worlds are not pertinent for present purposes, apart from the fact that laws of nature have a central position in fixing these truthvalues (and the laws that obtain in a given world again supervene on the distribution of the fundamental physical properties in the world in question).

For present purposes, the only point that we need is that the situation is not equivalent as regards the laws that support the mentioned two counterfactuals: the counterfactual “If p_1 had not occurred, p_2 would not have occurred either” can with reason be taken to express a causal relation between p_1 and p_2 , because it is supported by strict laws – at least if one takes p_1 and p_2 to be placeholders for appropriate configurations of fundamental physical tokens, p_1 standing for the fundamental physical supervenience base of m_1 . The laws in which types of physical properties figure are strict laws, or at least comparatively strict laws. Realism about entities and causal powers of entities combined with anti-realism about laws in fundamental physics (e.g. Cartwright 1983) is obviously not an option in the context of Humeanism and can in any case not help to make the assumption of systematic overdetermination plausible, since its commitment to a production view of causation simply rules systematic overdetermination out. The laws of the special sciences, by contrast, are never strict – or, in any case, less strict than fundamental physical laws, if one wants to leave the question of there really being strict laws open. If there are laws in the special sciences at all, these are *ceteris paribus* laws that admit a lot of exceptions, which cannot be specified in the vocabulary of the special science in question. This observation applies in particular to psychophysical laws. This was the reason why Davidson’s anomalous monism (Davidson 1970) could not show that events cause anything insofar as they instantiate a mental property (see the papers in Heil and Mele 1993). A similar objection applies in the present context.

Strong supervenience ensures that both the counterfactual “If p_1 had not occurred, p_2 would not have occurred either” and the counterfactual “If m_1 had not occurred, p_2 would not have occurred either” are true. However, the mere fact of these two counterfactuals being true does not ensure that they express a causal relation between the events in question each. The first one does so in virtue of its being supported by laws on a Humean regularity theory of causation, more precisely a sophisticated regularity theory that draws on laws insofar as they are central to making certain counterfactuals true. By contrast, in the case of the second counterfactual, there are no laws available that support the link between m_1 and p_2 , at least not in the same way as there are laws that support the link between p_1 and p_2 . As things stand, given the truth-makers for the first counterfactual, the second one is true only in virtue of m_1 strongly supervening on p_1 .

However, such an asymmetry just is the motivation for refusing to abandon (4). Supposing systematic overdetermination, what distinguishes a situation in which m_1 and p_1 both cause p_2 from a situation in which only p_1 causes p_2 , with m_1 being epiphenomenal, while supervening on p_1 ? For a Humean, it cannot be a primitive fact whether or not there is a causal relation between two events. A Humean regularity theory, and be it a sophisticated one in terms of counterfactual dependence, crucially draws on regularities that have the status of laws. But as regards laws, the situation is asymmetrical: there are no laws that support a mental to physical counterfactual in the same way as there are laws that support a physical to physical counterfactual. It is precisely this fact that drives the concern of epiphenomenalism, put in the

Humean context. The Humeans cannot help themselves to the fact of m_1 strongly supervening on p_1 , for that fact alone does not counter the epiphenomenalism objection: the mere fact of a property strongly supervening on another property cannot constitute a sufficient reason for claiming that the supervenient property also causes – some – of the effects that the subvenient property causes so that these effects are overdetermined. Such a claim would amount to simply stipulating that properties which strongly supervene on other properties cannot be epiphenomenal. It would thereby throw the discussion back into the state in which it was before Kim published his 1998 book, repudiating earlier work in which he took mere supervenience to be sufficient to ensure the causal efficacy of mental properties.

One may counter the preceding considerations by changing the position in such a way that one proposes to privilege the mental to physical counterfactual. The reasoning is, in brief, the following: the mental to physical counterfactual “If m_1 had not occurred, p_2 would not have occurred either” is privileged in certain situations on the grounds that m_1 is an occurrence of a determinable property and p_2 an occurrence of a more determinate property and that in certain situations it is the determinable by contrast to the determinate property occurrence that is causally relevant (see notably Yablo 1992 and Shoemaker 2007, chapter 2). Note, however, that the issue is about causation, not causal explanation. Consequently, if one holds such a view, one is (a) committed to regarding the distinction between determinables and determinates as an ontological rather than a conceptual one (see the objection by Gillett and Rives 2005). Furthermore, (b) one has to endorse *token* multiple realization, maintaining that the token m_1 could have been realized by a token of another type than the type of p_1 in another possible world. Token multiple realization can be taken to be outright absurd (Polger and Shapiro 2009). At least, it commits one to haecceitism: in this case, there are two property tokens m_1 and m_2 of the same mental property type; in the world w_1 , m_1 is realized by a physical configuration of type P_1 , and m_2 is realized by a physical configuration of type P_2 . In the world w_2 , by contrast, it is m_1 that is realized by a physical configuration of type P_2 , and it is m_2 that is realized by a physical configuration of type P_1 . The only difference between w_1 and w_2 is a swap of m_1 and m_2 . Worlds hence have to be recognized as different whose only difference consists in a swap of individuals, without there being any qualitative difference between them. Haecceitism is therefore widely considered to be an implausible position (see notably Lewis 1986b, chapter 4.4). Last but not least, (c) by privileging the mental to physical counterfactual, one no longer pursues a solution to the problem of mental causation that consists in abandoning the principle of no systematic overdetermination (4), but one rejects in fact the principle of completeness (3), endorsing physical effects that have mental in distinction to physical causes, insofar as this position remains committed to the principle of distinction (1) (see McLaughlin 2007 as regards that consequence).

We thus get back to the traditional options to solve the problem of mental causation by dropping either (3) and endorsing some sort of dualist interactionism or by dropping (1) and endorsing some version of the psychophysical identity theory. Given the firm scientific grounds for (3) (see e.g. Papineau 2002, appendix), the conclusion of Kim (1998) stands, which always was also the position of Lewis himself from his 1966 paper on, despite the recent discussion about systematic overdetermination: if one endorses realism about mental causation, one has to attack (1) and come to terms with a version of the identity theory, whatever the best physical theories of the real world are, and whatever is one’s preferred stance in the metaphysics of causation.

References

- Bennett, K. (2003). Why the exclusion problem seems intractable, and how, just maybe, to tract it. *Noûs*, 37, 471–497.
- Cartwright, N. (1983). *How the laws of physics lie*. Oxford: Oxford University Press.
- Davidson, D. (1970). Mental events. In: Foster, L., & Swanson, J. W. (eds.): *Experience and theory*. Amherst: University of Massachusetts Press, pp. 79–101.
- Gillet, C., & Rives, B. (2005). The non-existence of determinables: or, a world of absolute determinates as default hypothesis. *Noûs*, 39, 483–504.
- Hall, N. (2004). Two concepts of causation. In: Collins, J., Hall, N., & Paul, L. A. (eds.): *Causation and counterfactuals*. Cambridge (Massachusetts): MIT Press, pp. 225–276.
- Harbecke, J. (2008). *Mental causation. Investigating the mind's powers in a natural world*. Frankfurt (Main): Ontos.
- Heil, J., & Mele, A. (eds.) 1993. *Mental causation*. Oxford: Oxford University Press.
- Kim, J. 1998. *Mind in a physical world. An essay on the mind-body problem and mental causation*. Cambridge (Massachusetts): MIT Press.
- Kim, J. (2007). Causation and mental causation. In: McLaughlin, B. P., & Cohen, J. (eds.): *Contemporary debates in philosophy of mind*. Oxford: Blackwell, pp. 227–242.
- Kroedel, T. (2008). Mental causation as multiple causation. *Philosophical Studies*, 139, 125–143.
- Ladyman, J. (2008). Structural realism and the relationship between the special sciences and physics. *Philosophy of Science*, 75, 744–755.
- Lewis, D. (1966). An argument for the identity theory. *Journal of Philosophy*, 63, 17–25.
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70, 556–567.
- Lewis, D. (1986a). *Philosophical papers. Volume 2*. Oxford: Oxford University Press.
- Lewis, D. (1986b). *On the plurality of worlds*. Oxford: Blackwell.
- Lewis, D. (2004). Causation as influence. In: Collins, J., Hall, N., & Paul, L. A. (eds.): *Causation and counterfactuals*. Cambridge (Massachusetts): MIT Press, pp. 75–106.
- Loewer, B. (2007a). Counterfactuals and the second law. In: Price, H., & Corry, R. (eds.): *Causation, physics, and the constitution of reality. Russell's republic revisited*. Oxford: Oxford University Press, pp. 293–326.
- Loewer, B. (2007b). Mental causation, or something near enough. In: McLaughlin, B. P., & Cohen, J. (eds.): *Contemporary debates in philosophy of mind*. Oxford: Blackwell, pp. 243–264.
- Marras, A. (2007). Kim's supervenience argument and nonreductive physicalism. *Erkenntnis*, 66, 305–327.
- McLaughlin, B. P. (2007). Mental causation and Shoemaker-realization. *Erkenntnis*, 67, 149–172.
- Papineau, D. (2002). *Thinking about consciousness*. Oxford: Oxford University Press.
- Polger, T., & Shapiro, L. (2009). Understanding the dimensions of realization. *Journal of Philosophy*, 106, 213–222.
- Shoemaker, S. (2007). *Physical realization*. Oxford: Oxford University Press.
- Yablo, S. (1992). Mental causation. *Philosophical Review*, 101, 245–280.