

**THE DEBATE BETWEEN CAUSAL REALISM
AND CAUSAL CONSTRUCTIVISM:
METAPHILOSOPHICAL REFLECTIONS¹**

Erik Weber

ABSTRACT

In this paper I discuss, from a metaphilosophical point of view, the debate between causal realism and causal constructivism. First, I argue that the debate, if it is couched in the general terms as it is traditionally done, rests on a false dilemma. Then I argue that the debate must be disentangled into several more specific debates in order to be interesting.

1. Introduction

In this paper I discuss, from a metaphilosophical point of view, the debate between causal realism and causal constructivism. This is the most general debate with respect to the metaphysical status of causation: it is about whether causation exists in the external world, or not. Causal realism can be characterised as follows:

Causation is something that occurs in an external reality. The existence of causal relations does not depend on the presence of minds, speakers, observers or the like.²

¹ I thank Jan Willem Wieland and an anonymous referee for their comments on earlier versions of this paper. The research for this paper was supported by the Fund for Scientific Research - Flanders through project nr. G.0158.05.

² The definitions I give here are based on the definitions of Price 2001, p. 106

The opposite position, *causal constructivism* is then characterised as follows:

Causation is something that does not occur in an external reality. The existence of causal relations depends on the presence of minds, speakers, observers or the like.

In his essay on causality in the *Handbook of Philosophical Logic*, Jon Williamson starts his discussion of the metaphysical status of causation with the *convenience thesis*:

It is convenient to represent the world in terms of cause and effect. (2007, p 107)

The argument for the thesis is this:

It is convenient to represent the world in terms of cause and effect because a causal representation, if correct, enables us to make successful causal inferences: it allows us to make correct predictions, correct diagnoses and successful strategic decisions. (p. 107)

To the convenience thesis, Williamson adds the *explanation thesis*:

Humans think in terms of cause and effect because of this convenience, not because there is something physical corresponding to cause which humans experience. (p. 107)

This thesis sounds constructivist, but Williamson intends it to be weaker. In his view, the convenience is a sufficient explanation for the existence of causal beliefs: we don't have to assume that causal relations exist in order to explain the existence of causal beliefs. The convenience explanation ensures that we are not forced to adopt real causal relations as the only explanation of causal beliefs. It leaves open two possibilities:

The Explanation thesis divorces causal beliefs from any physical, mind-independent notion of causality. While one might remain agnostic as to whether there are physical causal relationships, one might instead adopt an *anti-physical* position, claiming that in the

interests of ontological parsimony one should reject physical causality. I leave the selection of an appropriate stance here entirely open. (p. 107)

Williamson's line of reasoning shows that causal realism is not trivially true: the debate between causal realism and causal constructivism is a real debate which deserves the attention of metaphysicians.

In Section 2, I argue that the debate, if it is couched in the general terms as we have done till now, is based on a false dilemma. In Section 4, I will argue that the debate must be disentangled into several more specific debates in order to be interesting.

2. A false dilemma

2.1 Ned Hall has recently argued that causation, "understood as a relation between events", comes in at least two basic and fundamentally different varieties:

Events can stand in one kind of causal relation - *dependence* - for the explication of which the counterfactual analysis is perfectly suited [namely, had *c* not occurred, *e* would not have occurred] (...). And they can stand in an entirely different kind of causal relation - *production* - which requires an entirely different kind of causal analysis [namely, *c* produces *e*] (...). (2004, p. 226, cf. pp. 252-257; emphasis added)

If we put his view into a definition of causation, we get:

C causes E, if and only if, [E counterfactually depends on C] or [there is a causal mechanism by which C produces E].

Hall argues that in some cases a reading in terms of production is required, in yet some others a reading in terms of dependence, and, in most cases, a reading in terms of both. Let us start with a quintessential example from Hall, which is intended to give an example of a case where only dependence is required:

Suzy and Billy have grown up, just in time to get involved in World War III. Suzy is piloting a bomber on a mission to blow up an enemy target, and Billy is piloting a fighter as her lone escort. Along comes an enemy fighter plane, piloted by Enemy. Sharp-eyed Billy spots Enemy, zooms in, pulls the trigger, and Enemy's plane goes down in flames. Suzy's mission is undisturbed, and the bombing takes place as planned. If Billy hadn't pulled the trigger, Enemy would have eluded him and shot down Suzy, and the bombing would not have happened. (p. 241)

Billy's pulling the trigger *did not produce* the bombing, rather it neutralized a state-of-affairs that would have prevented the effect from occurring. The occurrence of the bombing *was dependent* on Billy's pulling the trigger, but not produced by it. In this example, the effect counterfactually depends on the cause, but there is no mechanism linking cause and effect. Counterfactual dependence "seems to be the only appropriate causal relation for such "negative events" to stand in" (2004, p. 256). Hall counters the obvious counter-response that would read this example in mechanistic terms, as follows:

A remarkably frequent but entirely unsatisfactory response is the following: Billy's action *is* connected to the bombing via a spatiotemporally continuous causal chain – it's just that this chain consists, in part, of *omissions* (namely, the various failures of Enemy to do what he would have done, had Billy not fired). (...) For there is no reason to believe that the region of spacetime these omissions occupy intersects the region of spacetime that Suzy and her bomber *actually* occupy; to hold otherwise is just to mistake *this* region with the region she *would have* occupied, had Billy not fired. (p. 243)

In other cases there is a causal mechanism but no counterfactual dependence. Suppose that Billy and Suzy are engaged in a competition to see who can shatter a target bottle first. Suppose further that Suzy throws her rock a split second before Billy. Suzy's throw is spatiotemporally connected to the shattering in the right way, but Billy's is not:

Suzy's throw is a cause of the shattering, but Billy's is not. Indeed, every one of the events that constitute the trajectory of Suzy's rock on its way to the bottle is a cause of the shattering. But the

shattering depends on none of these events, since had they not occurred the bottle would have scattered anyway, thanks to Billy's expert throw. (p. 235)

This example shows that counterfactual dependence is only one of two possible forms of causation: there are cases in which there is causation but no counterfactual dependence..

An important aspect of Hall's view is that in typical cases of causation, both relations are present and production and dependence coincide (2004, p. 254 and p. 265). So though the relations are conceptually distinct, in the actual world their extensions overlap in most cases. Atypical cases of causation occur where there is a production relation without counterfactual dependence (e.g. overdetermination cases like the rock throwing) or where there is a relation of counterfactual dependence without a production relation (e.g. double prevention cases like Suzy's bombing raid).

2.2 Hall's dualistic theory of causation can be given a realistic interpretation. Such realistic interpretation consists in (i) the claim that the production relation is a real relation in the world, and (ii) the claim that the counterfactual dependence relation is a real relation in the world. Such a realist interpretation leads to a metaphysical puzzle: why are the extensions of these two metaphysically real but different relations approximately the same? The most plausible answer is that there is an unknown relation – the causal relation – which has the following properties:

- (1) The presence of the causal relation usually (but not always) leads to the presence of a production relation. Its absence usually (but not always) leads to the absence of a production relation.
- (2) The presence of the causal relation usually (but not always) leads to the presence of a counterfactual dependence relation. Its absence usually (but not always) leads to the absence of a counterfactual dependence relation.

This interpretation is by no means the only one. It is also possible to regard the production relation as a real relation in the world, and counterfactual dependence as a convenient human construction, or the other way around. Or one can regard both relations as human

constructions. Let us look at a philosopher – Phil Dowe – who chooses the first option.

2.3 Process theories of causation (e.g. Dowe 2000 and Dowe 2004) posit some sort of physical or material connection between causes and their effects. These theories experience difficulties when applied to causation by omission (cases where something that does not happen causes something) or prevention (cases where something causes something not to happen). Because of the absence of a physical or material connection in these cases, process theories fail to regard them as instances of causation, which leads to a number of judgments that are very counterintuitive, and, maybe even worse, seem to contradict science at some points. Consider the following examples.³ If we stick to a process theory of causation none of these propositions can be true:

- (1) A student's laziness (not studying enough) is the cause of his failing an examination.
- (2) The fact that I forgot to close the refrigerator is the cause of my ice cream melting.
- (3) The absence of vitamine C causes a person to have scurvy. (from Schaffer 2004, p. 202)
- (4) The failure of delivering a piece of machinery in time causes a machine to break down. (from Schaffer 2004, p. 201)
- (5) The fact that I did not give my pet plant any water caused its death.
- (6) A father pushing his child out of the way of a speeding car saved the child's life (i.e. causes it not to die).
- (7) A father not pushing his child out of the way of a speeding car is the cause of its death.
- (8) I caused the terrorist attack in London by failing to report information that I had about it. (from Dowe 2004, p. 191)

Counterfactual or counterfactual-based theories about causation have the opposite problem. They are forced to regard a large number of causal

³ These examples are taken from Froeyman & Weber (ms.)

propositions as true, despite the fact that they are intuitively false. For example:

- (9) My not throwing a rock causes the window not to break. (from Dowe 2004, p. 191)
- (10) You not throwing a rock causes the window not to break. (from Dowe 2004, p. 189)
- (11) I caused the terrorist attack in London by failing to be in a pub where I could have overheard the plot, and by failing to travel to the UK and blowing up the terrorist's van. (from Dowe 2004, p. 191)
- (12) The fact that you did not give my pet plant any water is the cause of its death.
- (13) The fact that it did not rain and that my roof did not leak is the cause of my plant's death.

The problems of the counterfactual approach can be easily solved if we acknowledge that our ordinary causal talk is determined by considerations of accuracy (these relate to the relation of our causal claims with the world) and considerations of adequacy (these relate to the reason our purpose we have when uttering a causal claim). This idea is certainly not new. We find it e.g. in John Mackie's *The Cement of the Universe*. He says that, if something is a cause in his sense (INUS condition plus causal priority) it is possible that:

... mention of this fact happens to be irrelevant, or less relevant than mention of some other cause of the same result, to some current purpose. (1974, p. 36)

If we apply the idea of adequacy relative to the purposes of causal claims, the problems mentioned above can be solved. The difference between (8) and (11) for example consists in the fact that, if you have information about an upcoming terrorist attack, we usually expect you to make it public. The fact that you did not make this information public, is important in the context of determining legal, moral and/or social responsibility. We do not expect you to know in which pub the plot is being made, and we certainly cannot expect you to be accidentally in this pub at the right moment. These facts are not important in the context of

determining responsibility. In the same sense, not throwing a rock is usually not considered a cause of the not breaking of a window, because people are usually not expected to throw rocks at windows (unless they suffer from a sort of compulsive rock-throwing obsession). If you own a plant, you are expected to give it enough water (as well as knowing more or less how much water it needs).

In his 2004 paper, Phil Dowe proposes to call intuitively acceptable cases of prevention and omission *quasi-causation*, as opposed to genuine causation, which presupposes a physical connection between cause and effect. This pair of concepts allows us to preserve a metaphysical distinction, while the overarching generic concept (“cause” without qualification) accounts for our ordinary causal talk.

2.4 Let us compare Dowe with Hall. For Dowe, the core of Hall’s dualism (the definition “C causes E, if and only if, [E counterfactually depends on C] or [there is a causal mechanism by which C produces E]”) is acceptable as a conceptual analysis of ordinary causal talk. However, he will say that only the first relation refers to something in the real world. His metaphysical position is an intermediate one:

Some types of causation occur in an external reality. The existence of causal relations of these types does not depend on the presence of minds, speakers, observers or the like. Other types of causation do not occur in an external reality.

This is a third possible position (besides the two positions mentioned in the introduction), which cannot be neglected. If one neglects this possibility, the debate is ill-constructed, because it rests on a false dilemma.

3. A range of more specific debates

3.1 The intermediate position described in 2.4 is very vague. To start with, it can cover Dowe’s position, but also the converse (where counterfactual relations are seen as real, while process causation is seen as a fiction). Furthermore, causation may involve more than two

relations. Hall himself indicates that dualism might not be enough. He gives the following counterexample to his own theory:

[T]here are certain kinds of cases that we have some inclination to call cases of causation, but that also elude classification in terms of production or dependence. Here is an example, a slight variation on the story of Billy, Suzy and Enemy: This time, there is a second fighter plane escorting Suzy. Billy shoots down Enemy exactly as before, but if he hadn't, the second escort would have. (2004, p. 271)

This example is an instance of preempted double prevention; in this case it is no longer true that the bombing wouldn't have happened if Billy hadn't pulled the trigger. Hall admits that one will nonetheless be inclined to grant Billy some causal responsibility for the success of the bombing, just as when there was no second escort. Cases like Hall's example with the second escort can be dealt with by means of a probabilistic theory of causation. So we could add a third relation, viz. unconditional probabilistic dependence to Hall's definition. That results in the following:

C causes E, if and only if, [E counterfactually depends on C] or [there is a causal mechanism by which C produces E] or [there is an unconditional probabilistic dependence relation between C and E].

As soon as we add a third relation in the conceptual analysis, the number of possible metaphysical positions that qualify as intermediate increases from two (Dowe's position and its converse) to six. If there happen to be more distinct causal relations, the number of positions increases in an obvious way.

Given the possibility of six intermediate positions and two extreme ones, the debate the metaphysical status of causation should be split up into the following three debates:

- (1) Does causation in the counterfactual sense really exist?
- (2) Does causation in the process sense really exist?
- (3) Does causation in the sense of unconditional probabilistic dependence really exist?

One's position in these more specific debates automatically determines one's general metaphysical position with respect to causation.

3.2 In Tooley (1986) a singularist account of causation is characterised as a positive answer to the following question:

[I]s it possible for two events to be causally related, without the relationship being an instance of some causal law, either basic or derived, an either probabilistic or non-probabilistic (p. 271)

Note that this question presupposes that there are causal relations at two levels: causal laws and causal relations between particular events. The question is whether relations of the last type can occur in cases where there is no relation of the first type. Both singularists and their opponents (the adherents of the so-called supervenience view), are realists about causation in a double sense: causal laws and singular causation. This double realism is a presupposition of their debate.

However, this presupposition can be challenged. Suppose that I accept that type-level causation is adequately characterised by a probabilistic conception. Take for instance the following definitions, which constitute the core of Ronald Giere's theory of probabilistic causation in populations:

C is a *positive causal factor* for *E* in the population *U* whenever $P_X(E)$ is greater than $P_K(E)$.
C is a *negative causal factor* for *E* in the population *U* whenever $P_X(E)$ is less than $P_K(E)$.
C is *causally irrelevant* for *E* in the population *U* whenever $P_X(E)$ is equal to $P_K(E)$. (Giere 1997, p. 204)

Though it can be extended to other types of variables, Giere considers only binary variables. So in his definitions, *C* is a variable with two values (*C* and not-*C*); the same for *E* (values *E* and not-*E*). *X* is the hypothetical population which is obtained by changing, for every member of *U* that exhibits the value not-*C*, the value into *C*. *K* is the analogous hypothetical population in which all individuals that exhibit *C* are changed into not-*C*. $P_X(E)$ and $P_K(E)$ are the probability of *E* in respectively *X* and *K*.

An example might clarify this. If we claim that smoking (**C**) is a positive causal factor for lung cancer (**E**) in the Belgian population (**U**), this amounts to claiming that if every inhabitant of Belgium were forced to smoke there would be more lung cancer patients in Belgium than if everyone were forbidden to smoke. Conversely for the claim that smoking is a negative causal factor. Causal irrelevance is a relation between variables (represented in bold) rather than a relation between values of a variable (like the first two relations). If we claim that “smoking behaviour” (**C**) is causally irrelevant for “the occurrence or absence of lung cancer” (**E**) this means that we believe that in the two hypothetical populations the incidence of lung cancer is equally high.

It is perfectly possible to accept that this theory captures the way causation is used in the sciences, and at the same time be an anti-realist about these causal relations. In that case, we regard them as useful fictions (e.g. useful for policy). A possible argument for such a constructivist position could be that the definitions refer to hypothetical worlds. So one has to be a modal realist in order to accept them as descriptions of a real relations.

Such constructivism at the type-level can be combined with realism at the singular level, for instance by accepting the view that singular causation is causal mechanical (and claiming that these are real relations in the world, see Section 3.1). Given the coherence of this intermediate position, it is better to split up the debate on causal realism in two subordinate debates, dealing with the following questions:

- (1') Does singular causation really exist?
- (2') Are there causal laws in the real world?

4. Conclusion

I have argued that, in order to avoid a false dilemma, we have to take into account an intermediate position in the debate between causal realists and casual constructivists:

Some types of causation occur in an external reality. The existence of causal relations of these types does not depend on the presence of minds, speakers, observers or the like. Other types of causation do not occur in an external reality.

I have also shown that the debate is better split up into several smaller debates, along two lines. On the one hand, we have subdebates based on the type of causal relation::

- (1) Does causation in the counterfactual sense really exist?
- (2) Does causation in the process sense really exist?
- (3) Does causation in the sense of unconditional probabilistic dependence really exist?

On the other hand, we have subdebates based on the level of causation:

- (1') Does singular causation really exist?
- (2') Are there causal laws in the real world?

There are important open questions about how do these two strands of subdebates relate to each other. Can they be treated independently of one another? Does it matter where we start? Do we have to make further subdivisions by combining the two strands?

Ghent University (UGent)
Email: Erik.Weber@UGent.be

REFERENCES

- Dowe Phil (2000), *Physical Causation*. Cambridge: Cambridge University Press.
- Dowe Phil (2004), 'Causes Are Physically Connected to Their Effects', in C. Hitchcock (ed.), *Contemporary Debates in Philosophy of Science*. Oxford: Blackwell, pp. 189-196.
- Froeyman Anton & Weber Erik (ms), 'Prevention and Causation by Omission: Where Is the Metaphysical Problem?'
- Giere Ronald (1997), *Understanding Scientific Reasoning* (4th edition). Fort Worth: Harcourt Brace College Publishers.
- Hall Ned (2004), 'Two Concepts of Causation', in J. Collins, N. Hall & L.A. Paul (eds.), *Causation and Counterfactuals*. Cambridge, MA: The MIT Press, pp. 225-276.
- Mackie John (1974), *The Cement of the Universe. A Study of Causation*. Oxford: Clarendon Press.
- Price Huw (2001), 'Causation in the Special Sciences: the Case for Pragmatism', in M. C. Galavotti, P. Suppes & D. Costantini (eds.), *Stochastic Causality*. Stanford: CSLI Publications, pp. 103-121.
- Schaffer Jonathan (2004), 'Causes Need not to Be Physically Connected to Their Effects', in C. Hitchcock (ed.), *Contemporary Debates in Philosophy of Science*. Oxford: Blackwell, pp. 198-216.

- Tooley Michael (1986), 'The Nature of Causation: a Singularist Account', in *Canadian Journal of Philosophy* **16**, pp. 271-322.
- Williamson Jon (2007), 'Causality', in D. Gabbay & F. Guentner (eds.), *Handbook of Philosophical Logic, volume 14*. Dordrecht: Springer, pp. 95-126.