



Universiteit  
Leiden  
The Netherlands

## **How (not) to judge a theory of causation**

Gijsbers, V.A.

### **Citation**

Gijsbers, V. A. (2021). How (not) to judge a theory of causation. *Synthese*, 199, 3117-3135. doi:10.1007/s11229-020-02927-2

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/3239289>

**Note:** To cite this publication please use the final published version (if applicable).



# How (not) to judge a theory of causation

Victor Gijsbers<sup>1</sup>

Received: 3 May 2019 / Accepted: 20 October 2020  
© The Author(s) 2020

## Abstract

Philosophical theories of causation are commonly judged by their ability to correctly determine whether there is a causal relation present in intuitively clear example scenarios. If the theories survive this test, they are then used to answer big philosophical questions about causation. This Method of Examples is attractive because it seems to allow us to determine the quality of a theory of causation independently of answering the big philosophical questions; which is good, since it means that we can then non-circularly use the theories judged to be best to answer those questions. However, the current article argues that this virtue of the Method of Examples is an illusion. In particular, I argue that the necessary step of judging whether a proposed analysis of causation is reductive can only be taken after many of the most vexing philosophical questions about causation have already been answered. It is then shown that a rejection of the methodological supremacy of the Method of Examples leads to a more pluralistic method of judging theories of causation, a pluralism that benefits non-standard approaches like interventionism and agency theories.

**Keywords** Causation · Counterexamples · Reductive analysis · Interventionism · Agency theories of causation

## 1 Introduction

Philosophical discussion of causation has bloomed during the past decades. We have come a long way since Bertrand Russell (1913) and Rudolf Carnap (1928) argued that causation is nothing but the kind of functional relation encapsulated in mathematical laws; and friends of causation can now choose from a wide variety of detailed and sophisticated theories. There are regularity theories, counterfactual theories, causal process theories, probabilistic theories, interventionist theories, and more. The richness

---

✉ Victor Gijsbers  
V.Gijsbers@hum.leidenuniv.nl

<sup>1</sup> Universiteit Leiden, Institute for Philosophy, P.O. Box 9515, 2300 RA Leiden, The Netherlands

of the field is almost an embarrassment, given that most of these approaches are mutually incompatible.

As is to be expected, much attention has been given to *judging* theories of causation: finding out which ones are better and which ones are worse. Often, this is done by confronting the theories with elaborate examples and counterexamples. Indeed, the literature on causation is somewhat notorious for the fanciful scenarios it contemplates, ranging all the way from simultaneous stone throwings and evil backup assassins to unreliable sorcerers casting identical spells and the queen of England not stopping by to water my plants. Clearly, the idea is that a theory of causation that is worth its salt should be able to make correct pronouncements—both in normal and in abnormal situations—about which events cause which other events.

On second thoughts, however, this desideratum may seem rather strange. Having a theory that makes such pronouncements might be of paramount importance if we want to implement causal reasoning in AI, or develop a statistical method for discovering cause-effect relations. And some philosophers are indeed engaged in such projects. But one would expect that the *primary* goal philosophers try to achieve with a philosophical theory of causation is to answer some of the Big Philosophical Questions (BPQs) about causation, questions such as:

Is causation an objective relation? Do causal facts supervene on non-causal facts?

Is causation a fundamental metaphysical ingredient of our world? What is the relation between the direction of causation and the direction of time? What is the relation between causation and agency? How do we learn about causal facts?

Can causation be perceived?

Note that none of these questions concern assassins or the queen of England. So why is it the case that philosophers are often—though by no means exclusively—focused on assessing theories through examples and counterexamples, given that the examples and counterexamples are not themselves what exercises the philosophical imagination? The answer, as I will argue at some length in the next section, is this: judging theories based on examples appears to give us a *non-circular* way of answering the BPQs, because it allows us to assess the correctness of theories of causation without presupposing that we already know the answers to the BPQs.

Given that non-circularity is a methodological virtue, there is thus a methodological presumption in favour of what we can call the *Method of Examples*. However, this article's main thesis is that this presumption, while influential in the literature, is nevertheless false. It is false because (a) the presumption requires that the theories assessed by the Method of Examples are *reductive analyses*, that is, that they analyse the notion of causation in non-causal terms; but (b) assessing whether or not an analysis is in fact reductive requires us to first answer many of the BPQs. Hence, the Method of Examples is after all circular: it has to presuppose answers to the BPQs in order to assess whether a theory of causation gives us the right answers to the BPQs. This does not mean that the Method of Examples is worthless; but it does mean that it has no special authority. This will be shown to have important consequences for the philosophy of causation.

The article is structured as follows. In Sect. 2, I flesh out the schematic claims made above by discussing the methodological presumption in favour of the Method of

Examples, and showing why this Method requires theories of causation to be reductive analyses. Then, in Sects. 3 to 5, I attack the presumption of non-circularity by arguing that answers to important BPQs are presupposed in judging whether or not a theory of causation is reductive. Specifically, I will look at the relation between causation and properties in Sect. 3; at the relation between causation, worlds and reality in Sect. 4; and at the relation between causation and spatio-temporal relations in Sect. 5.

Before I move to establish some consequences of this attack, Sect. 6 is a brief intermezzo to clarify the relation between my approach and that of two related trends in the philosophy of causation: first, the anti-reductionism of Woodward (1990) and Carroll (1994, 2009); and second, the attack on the use of examples and counterexamples by Glymour et al. (2010). Finally, in Sect. 7, I discuss the methodological lessons to be drawn from the paper, and illustrate these lessons with two philosophical approaches to causation: interventionism (Woodward 2003) and agency theories of causation (Collingwood 1940, Von Wright 1971, Menzies and Price 1993). Both of these approaches will come out looking better than they did after the lessons have been applied.

## 2 Examples and reduction

The aim of a philosophical theory of causation can be supposed to be, primarily, answering the BPQs about causation. It may seem, then, that it is also clear by which standard such theories must be judged: a theory is *good* if and only if it gives the *correct* answers to the BPQs.

But when we try to apply this standard, we run head-first into Meno's paradox (Plato 1997, *Meno*, 80d). In order to assess whether a theory of causation gives us the correct answers *are*; but if we already know what the correct answers are, we are no longer in need of a theory of causation. On the other hand, if we *don't* know what the correct answers are, then it seems that we are unable to judge theories of causation at all; this means we can only adopt one at random without knowing whether it is any good, hence, without being able to trust the answers it gives us. As Socrates expresses it: "He cannot search for what he knows—since he knows it, there is no need to search—for what he does not know, for he does not know what to look for." (Plato 1997, *Meno*, 80e).

Of course, this is a perfectly general problem, not limited to theories of causation; and there are many general solutions on offer. We can, for instance, judge our theories in terms of their ability to (1) do justice to our pre-theoretical commitments; (2) achieve reflective equilibrium with our other philosophical and scientific beliefs; or (3) be derived by a priori methods. In the philosophy of causation, a particular version of (1) has become especially prominent.

When philosophers start discussing causation, they quickly realise that they can all agree on certain examples in which there either is or is not a causal relation between two events. If you and I both throw a brick at a window, and my brick hits the window first, while yours passes through the air where the window used to be only a second ago, then we can surely agree that my throwing caused the window's breaking and yours did not.

Our agreement about this scenario is independent of our more ‘philosophical’ beliefs. We may have no idea how to answer questions about the metaphysics or epistemology of causation, and yet we know perfectly well which events do and do not cause which other events in the window scenario. We can apparently agree about the *causal facts* while disagreeing or being ignorant about the *nature of causation* and the answers to the BPQs.

This suggests an easy and intriguing way out of the Meno paradox. Suppose that we are confronted by some theories, each claiming to tell us about the nature of causation. Now we cannot judge these theories based on whether they are right about the nature of causation, for the reason given by Meno and Socrates. But what we *can* do is judge these theories on their ability to *save the causal facts*, that is, make true pronouncements about which events do and do not cause which other events; and then we make the assumption—surely plausible?—that if a theory manages to do this, it might well be true, and hence a good candidate for helping us understand the nature of causation.

A single glance at the literature on causation suggests that this way out of the Meno paradox is almost universally embraced and applied: everywhere we see theories of causation being judged based on examples and counterexamples at the level of causal facts. (Other criteria for assessment are also used, but less universally.) Let us call this way out of the paradox the *Method of Examples*, since it is a method that consists of first judging causal theories based on finding out whether the theory gets specific *examples* of cause-effect relations right, and that only then applies successful theories to the BPQs.

Compared to other methods for judging theories of causation, it seems to have a clear advantage: it is non-circular. For instance, assessing a theory of causation in terms of its reflective equilibrium with our other philosophical theories will generate circularities: if I am a Humean, Humean theories look good; if I’m an anti-Humean, then Humean theories look bad. The Method of Examples offers a way out of this conundrum by promising non-circularity and building only on common ground—hence, there is a methodological presumption in favour of using this Method. Therefore, any theory of causation that looks bad when judged by the Method of Examples will have a tough time persuading philosophers of its qualities, even if it looks better when assessed by other standards.

Now in order for the Method of Examples to be able to applied to it, a theory of causation must meet three requirements. First, the theory must give conditions for the truth of causal claims, that is, it must have roughly this form:

$c$  is a cause of  $e$  if and only if ...

since it is precisely statements of this form that determine the extension of the ‘is a cause of’ relation, and which can therefore be tested using examples and counterexamples. Indeed, the major contemporary theories of causation all have this form. I present a few in a simplified fashion:

1. *The regularity theory of Hume (1748)*:  $c$  is a cause of  $e$  if and only if  $e$  happens just after and very close to  $c$  and events of type  $C$  are always followed by events of type  $E$ .

2. *The probability raising theory of Reichenbach (1956)*  $c$  is a cause of  $e$  if and only if  $P(eltc) > P(elt\text{not-}c)$  and there is no earlier event  $b$  that screens  $e$  off from  $c$ .
3. *The counterfactual theory of Lewis (1973)*  $c$  is a cause of  $e$  if and only if, if  $c$  were to occur  $e$  would occur; and if  $c$  were not to occur  $e$  would not occur.
4. *The interventionist theory of Woodward (2003)*  $c$  is a cause of  $e$  if and only if there is a possible intervention on  $c$  such that it changes  $e$  or the probability of  $e$ .<sup>1</sup>

These theories all have the ‘canonical’ form of giving conditions that are both necessary and sufficient, though it is of course also possible to test theories that give only sufficient, or only necessary conditions; or even some combination of sufficient conditions and necessary conditions that leave a certain space within which causal claims are not assigned a determinate truth value.<sup>2</sup> But for present purposes we can focus on theories with the canonical form, for reasons that will become clear very soon.

That a theory has the right form makes it possible to check whether it saves the causal facts, which is the first step in the Method of Examples; but the crucial second step of the Method consists in then using the theory to answer philosophical questions about causation. The procedure for doing this is as follows. We take a BPQ; we answer it for the right-hand side of the theory; and then we identify the answer for the right-hand side with the answer for the left-hand side. For example, we take the regularity theory of Hume and the question whether causation is an objective relation. What appears on the right-hand side of the theory, spatio-temporal contiguity and regular succession, certainly are objective relations. So we conclude that causation is an objective relation too.

But in general this procedure is only possible if the theory in question expresses an *identity*, that is, if (a) the theory is indeed in canonical if-and-only-if form; and (b) this “if and only if” should be read as “is nothing over and above,” that is, it should be read as establishing an identity of meaning that guarantees substitutability in intensional contexts. Without the criterion of identity, there is no guarantee that answers about the right-hand side carry over to the left-hand side; and without that guarantee we would not be justified in using the theory to answer BPQs.<sup>3</sup>

Furthermore, it must be the case that the right-hand side of the theory contains no causal terms. In other words, the theory must be *reductive*. For without reduction, we would enter a circle in which we need to have the answers to our philosophical questions about causation before we can use the theory to arrive at those answers. In

<sup>1</sup> The first three theories are not precisely recent, but many recent theories are sophisticated developments of these originals. For regularity theories, see Davidson (1967), Mackie (1974), Baumgartner (2008), Paul and Hall (2013); for probabilistic theories, see Suppes (1970), Cartwright (1979), Kvart (2004), Hitchcock (2016); for counterfactual theories, see Lewis (2000), Hitchcock (2001), Paul & Hall (2013), Menzies (2014). Note that this three-fold classification is somewhat loose and that many theories combine features of two or more approaches. In the case of Woodward’s theory, I have been guilty of significant changes—his theory is actually stated in terms of *direct* causation between *type-level* causes relative to a variable set  $V$ . This is very important for an understanding of interventionism, but can be ignored for present purposes.

<sup>2</sup> I’m indebted to an anonymous referee for insisting on this point.

<sup>3</sup> Compare: “ $X$  is a beautiful painting only if  $X$  is a painting. Calling something a painting does not involve an aesthetic judgment. So calling something a beautiful painting does not involve an aesthetic judgment.” Valid inference from the right-hand side to the left-hand side requires, in general, an identity of the two sides, which is clearly missing if we supply only necessary or only sufficient conditions.

the Humean example above, if spatio-temporal contiguity were defined in terms of causation, then we could not know whether it was an objective relation until we knew whether causation was an objective relation—and hence we would be engaged in a circle.

Here are two toy theories that exhibit the two failings of not positing an identity and of not being reductive:

- $c$  is a cause of  $e$  if and only if God believes that  $c$  is a cause of  $e$ .
- $c$  is a cause of  $e$  if and only if  $e$  is an effect of  $c$ .

The first theory, given that it is not to be read as an identity but merely as a partial expression of God's omniscience, doesn't tell us anything about causation; for instance, it does not allow us to conclude that causal relations are due to God's beliefs. The second theory is equally unhelpful, since any philosophical question we might have about causes is surely a question we also have about effects. Moving from toy theories to real ones, we notice that theories 1 to 3 are all explicitly meant as reductive analyses. Number 4, interventionism, is meant to express an identity but *not* to be reductive; and it is often criticised for precisely this reason (see Gijsbers and De Bruin (2014) for an overview of such criticisms).<sup>4</sup>

That a theory is meant to express an identity can simply be postulated when the theory is presented. But that it is reductive is a claim that must be judged on its merits. In particular, we must take a good hard look at the theory's right-hand side to see whether or not it contains any causal terms; any terms that either explicitly or implicitly call on the very idea of causation that we are attempting to analyse. What is more, this judgement must be made independently of our answers to the BPQs. Since the methodological presumption in favour of the Method of Examples is based precisely on its promise to give us a non-circular way of answering the BPQs; and since being reductive is a requirement for a theory to be able to fulfil this purpose; therefore the methodological superiority of the Method of Examples would be revealed as an illusion if it turned out that answers to the BPQs *are* presupposed in judgements about the reductive nature of a theory.

As was indicated in the introduction, the main aim of the current article is showing that the methodological superiority of the Method of Examples is indeed an illusion, for the reason indicated. Showing this is the task of the next three sections.

### 3 Causation and properties

This section and the next two will attempt to establish the thesis that judging whether or not a specific analysis of causation is reductive requires us to *first* answer important and controversial questions about causation. An analysis is reductive if and only if its right-hand side contains no terms that overtly or covertly appeal to the notion of causation. Let us call such terms 'causally innocent' and all other terms 'causally culpable'.

Now it is absolutely crucial that we realise that to be a reductive analysis, it is not just the terms that *overtly* appear on the right-hand side of the theory that must be causally

<sup>4</sup> We will discuss the status of interventionism at more length in Sect. 7.

innocent, but also the terms presupposed in the application of the right-hand side criterion. Take, for example, the counterfactual theory of Lewis. In order to judge a causal claim, it asks us to evaluate a counterfactual. Evaluating a counterfactual requires us to take a full description of a possible world, that is, a full specification of how events are distributed across a particular spacetime manifold, and then assess ‘distances’ between this possible world and other possible worlds. If any of this machinery requires the use of causally culpable terms, then the counterfactual theory is not reductive.

These three sections will focus on three types of ingredient that *any* analysis of causation is likely to need. First, we look at the properties needed to describe the content of a world; second, we look at the idea of *a world* or *the world* itself; and finally, we look at the spatiotemporal framework within which events are to be described. Again, the aim is not to prove that these ingredients cannot be causally innocent. Rather, it is to show that to claim that they *are* causally innocent is to presuppose answers to important BPQs.

Let us start with the terms that we need to describe the contents of a world; the very events that appear as *e* and *c* in all our four example theories. The question is then whether the set of causally innocent terms is rich enough to describe what happens in a world. The idea that describing a single event requires the use of causally culpable terms may seem far-fetched, but in fact most descriptions of events are clearly causal in nature. “An elephant is in the room” *my seem* causally innocent. In fact, however, to be an elephant, something must be related to paradigmatic examples of elephants through a series of “offspring of” and “parent of” relations—and these are evidently causal terms. (To be something’s parent is, in part, to have caused that thing to exist.)

Now the concept of elephant may not be absolutely necessary for describing the world. It would perhaps be more worrying, especially to philosophers of an empiricist bent, if the basic concepts of human perception turn out to be causally culpable. But, as Keith Allen points out:

[c]ontemporary discussions of colour often start from the intuition that, whatever else they might be, colours are those properties that at least normally *cause* colour experiences. (Allen 2016, p. 96; emphasis added.)

Allen goes on to (critically) discuss some non-causal theories of colour, but especially interesting is his argument that resistance to causal theories of colour is often based on prior acceptance of the Humean principle that causes cannot logically imply anything about their effects. If being red *is* causing red experiences, then the cause and the effect do seem to be logically linked. Allen suggests that we can simply deny the Humean strictures on causation (p. 100). For our purposes, it does not matter whether the Humean is right or wrong; what matters is that we apparently need to already embrace Humean answers to the BPQs before we can argue that perceptual properties are causally innocent.

But perhaps we should not focus on perception, since contemporary analytic philosophers are likely to argue that the most basic description of the world is not in terms of perceptual properties, but rather in terms of fundamental physics. However, here too doubts about causal innocence can be raised. It is *prima facie* plausible that to have positive electrical charge *is* (at least in part) to attract negatively charged objects and repel other positively charged objects with a force given by Coulomb’s

law. And ‘attract’ and ‘repel’ are obviously causal notions: to attract something is to cause it to move towards the attracting body; to repel something is to cause it to move away.

This analysis of charge can no doubt be resisted; and it could also be argued that charge is not, perhaps, one of the *fundamental* ingredients of the ontology of the world. So the crucial question is whether the fundamental ontology of the world will or will not involve causation. This is a hotly contested issue in contemporary metaphysics. On the one hand we can find modern-day Humeans like David Lewis, who contend that it will not:

all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. (Lewis 1986: ix)

On the other hand, there is David Armstrong’s (1983) theory of universals, which holds that causal relations are founded on a relation of natural necessitation; and, especially interesting when we are thinking about properties, a variety of recent proposals concerning the fundamentality of essences and dispositions. For instance, Ellis 2001 holds that being a certain kind of event—such as being a charged particle appearance—essentially involves standing in certain kinds of causal relations to other events. The emerging theory of *pandispositionalism* (Mumford 2004; Mumford and Anjum 2011; Bird 2007, 2016) even goes as far as to claim that all properties in nature are dispositions and that dispositional terms will therefore be an irreducible part of any description of the world. The concepts of disposition and cause are intimately related, which makes it likely that dispositional descriptions of the world will turn out to be causally culpable.

Again, we are not now interested in finding out which of these positions is correct. The point is that before we can assess whether an analysis of causation is reductive, we have to know whether it is possible to give a causally innocent description of the contents of the world. But this is precisely one of the most hotly debated BPQs about causation. Thus, at least one important BPQ has to be answered before we can start using the Method of Examples to answer BPQs.

## 4 Causation and worlds

An easily overlooked ingredient of analyses of causation is the notion of a world, or, in some case, of *the* world. The idea of *a* world comes up in all analyses that are meant to be applied to possible as well as actual scenarios, the case on which I will focus first. For example, according to the Humean theory, *c* causes *e* just in case events of type *C* are always followed by events of type *E*—but this condition is implicitly to be read as having to hold only *in the world that contains c and e*. This latter phrase is a crucial addition, since without it, the theory would require constant conjunction in every possible world, which is a condition that would never be met (at least not given a Humean metaphysics). For analogous reasons, Reichenbach’s probabilities are to be evaluated within a single world; and Lewis’s account of counterfactuals is completely based on the notion of a possible world.

There seems to be no way around the invocation of worlds. Suppose that we are given a set of two events: Jack ate a pound of arsenic at 8:00 pm; Jack died at 8:04 pm. Was Jack's death caused by arsenic poisoning? Any philosopher of causation well-trained in the Method of Examples knows that this depends on the occurrence or non-occurrence of other events. Was it, for instance, the case that an anvil fell on Jack's head at 8:04 pm? Was it the case that Jack ate an antidote to arsenic at 8:02 pm? Now suppose that we are given some list of events,  $L$ , and are asked to determine whether two of those events,  $A$  and  $B$ , are related as cause and effect. It is impossible to do that. We need something more: we need an assurance that  $L$  contains *all of the relevant events*, that is, all of the events we need to know about in order to answer the question correctly.

But how are we to define relevance? We could define it in terms of counterfactuals and state that we need  $L$  to be such that no events not in  $L$  made a difference to the occurrence of  $B$ . But then we need to be able to assess counterfactuals before we can assess a counterfactual, thus embarking on a regress. We could also say that we need  $L$  to be such that it describes a causally closed system, or that all events not in  $L$  had a causal influence on  $B$  only through events in  $L$ . But we were trying to define causation in terms of counterfactuals, and now we have to invoke causation; so this too leads to a regress. For these reasons, counterfactual theories of causation usually invoke a concept that seems causally innocent and is nevertheless strong enough to imply causal isolation, namely, the concept of world. Worlds are generally agreed to be causally isolated from each other; so if we are given a complete description of a world, it is guaranteed to list all the events we need to make correct causal judgements.

(But couldn't we make do with less than a world? For instance, given events  $c$  and  $e$ , might it not be enough—for some analyses, at least—to have a full description of the events in the intersection of  $c$ 's forward and  $e$ 's backward light cone? Perhaps; but *that* this is all we need is because the world has specific laws of nature, e.g., those of special relativity. An analysis that is supposed to be correct in all possible worlds cannot assume that particular laws of nature hold. Nor would it help to appeal to the concept of a law of nature, since an analysis of this concept would surely itself invoke the concept of world.)

So we must raise the following question: might the concept of world itself be a causal concept? Using Lewis's 1986 terminology, the question is whether the relation of *being worldmates*, which is the relation that holds between two possible individual just in case they belong to the same world, is causally innocent or causally culpable. Now we have already remarked that worlds are causally isolated from each other; and this is of course not an empirical discovery, but rather a necessary truth. This suggests that perhaps it is a necessary truth because it is an analytic truth about the concept of world: worlds just *are* maximal mereological sums of causally related individuals. To be the worldmate of  $X$  just is to stand in some—possibly infinite—series of causal relations to  $X$ .

Thus, Aristotle is my worldmate because there is a direct line of causal links from some event involving Aristotle (his writing of the *Metaphysics*) to some event involving me (my reading of the *Metaphysics*). But even a philosopher outside of my light cone can be my worldmate by virtue of a series of events  $e_1, e_2, \dots, e_n$ , such that I am causally

related to  $e_i$ , she is causally related to  $e_n$ , and there is a causal relation between  $e_i$  and  $e_{i+1}$  for all  $i$ .

If this is the correct way of understanding worlds, then invoking the concept of world is invoking the concept of causation. To claim that there is constant conjunction of  $C$ s and  $E$ s in the actual world would be to claim that all  $C$ s that are causally connected to us are followed by  $E$ s. This would radically undermine the claim that the proposed analysis of causation is reductive—and no analysis could escape from this predicament. This is also true for theories that strive merely to be correct in the *actual* world, that is, in *the* world; for such theories still need to require that they be given a full description of everything that is *real* or *actual*; and on the current proposal, this is simply to ask for all the events that are causally connected to us.<sup>5</sup>

Of course, it is possible to embrace an alternative theory of worlds. There is a prominent one on the books in analytic metaphysics,<sup>6</sup> namely Lewis's proposal that:

whenever two possible individuals are spatiotemporally related, they are worldmates. If there is any distance between them—be it great or small, spatial or temporal—they are parts of one single world. (Lewis 1986, p. 70)

According to Lewis, this is a sufficient and necessary condition: worlds just are maximal sums of spatio-temporally related things. Perhaps he is right. But he isn't *obviously* right, and his proposal needs to be defended against suggestions like the following:

1. There could be worlds that are not spatio-temporal. Lewis himself considers it a serious possibility that some worlds may be ordered by relations that are only analogous to spatio-temporal relations and he spends some time (1986, pp. 74–78) attempting to fit this into his theory. At the end of this passage, he calls his attempt “an underdeveloped suggestion.”
2. There could be worlds that consist of causally connected but spatio-temporally disconnected regions. For instance, imagine two planets that do not have a spatio-temporal relation to each other (no amount of travel could get us from one to the other; no series of light beams trace a path between them) but where the act of sneezing instantly moves the inhabitants from one planet to the other or back again. The burden would seem to be on the Lewisian to argue that such a scenario is metaphysically impossible. (The Lewisian cannot argue that in this case we have a causal connection between two different worlds, because then the crucial idea of the causal isolation of worlds—to which Lewis commits himself—would be lost.)
3. Even if spatio-temporal relatedness were a necessary and sufficient condition for being worldmates, this might be because (a) causal relatedness is a necessary and sufficient condition for being worldmates and (b) it is a necessary truth that causal relations can only exist across space and time. If so, the concept of world would still be a causally culpable concept.

<sup>5</sup> For the actual world, the proposed causal criterion of belonging to a world thus reduces to the so-called Eleatic Principle; see, e.g., Oddie (1982).

<sup>6</sup> The notion of world also has a prominent place in continental phenomenology, e.g., in Heidegger (1927). But these theories offer little comfort to someone who wants to defend the possibility of reductive analysis. It is hard to claim that we have reduced away causation if we ‘reduce’ it to an agent’s meaningful practical engagement with her life world.

Holding on to the idea that a reductive analysis is possible, then, requires us to defend a view of worlds (or *the* world) as not constituted by causation, a defence which involves taking a stance on a fundamental question about causation. So, again, we must answer a BPQ before we can use the Method of Examples to answer any BPQs.

## 5 Causation, space and time

To describe a world, we not only need to describe all the events in that particular world, but we also need to give the spatio-temporal relations between those events. Knowing that a hammer was swung and knowing that a vase broke doesn't teach us much unless we also know that the swinging happened right next to and just before the breaking. In order to assess whether reductive analyses of causation are possible, we must therefore also ask whether spatial and temporal concepts such as 'next to' and 'before' are causally innocent. In this section, I will ignore spatial relations and focus exclusively on temporal ones, since there is a richer tradition to draw on.

The most ambitious and radical attempts to argue that temporal relations are causally culpable were the so-called *causal theories of time*. While the idea can be traced back to Leibniz, it was especially in the 20th century and under the influence of Einstein's theory of special relativity that a number of authors—including Reichenbach 1956, Grünbaum 1968, Van Fraassen 1970 and Mehlberg 1980—attempted to analyse temporal relations in terms of causal relations. For instance, such a theory might claim that for  $a$  to be later than  $b$  just is for there to be a potential causal influence from  $b$  to  $a$ . (This theory is based on the idea that the forward light cone of an event  $a$  consists of all and only those events that can be potentially causally influenced by  $a$ , and that those same events are the ones that every inertial observer would agree to be in  $a$ 's future.)

Admittedly, these theories have fallen out of favour, at least partly because of powerful criticisms levelled against them; see, e.g., Smart 1969 and Earman 1972. And perhaps abandoning causal theories of time was well-justified—I will not try to adjudicate that issue here (but see my Gijsbers 2019). However, even if temporal relations cannot be *reduced* to causal relations, there still might exist an intimate connection between the two; a connection so intimate that understanding temporal concepts would necessarily involve understanding causal concepts. One philosopher who argued this is Kant, whose famous second analogy of experience (in Kant 1781/1787) leads to the conclusion that experience can only be temporal if the experienced events are thought under the category of causation. Rather than attempting to unravel Kant's notoriously difficult text and engaging with his elaborate metaphysical system, I will provide an argument of my own that is, let us say, *loosely* inspired by his.

Here are two events in the history of our world: ( $a$ ) you open this document; ( $b$ ) you remember opening this document. Of course,  $a$  happened before  $b$ . But how do you know this? What allows you to make a judgement about the temporal order of your experiences? This is a puzzling question, because time itself is not experienced: it is not part of the phenomenal content of an experience that it happens at a particular point in time. Seeing a blue square in 2019 is not a different experience from seeing a blue

square in 2020. In both cases the experience is one of *seeing a blue square*, not one of *seeing a blue square at time such-and-so*. Nor does the content of our memories change in a systematic way as time goes by. Older memories are not necessarily more vague, or less detailed, or more purple; nor are they characterised by any other phenomenal difference. So it is not possible to just compare the phenomenal contents of your two memories and see that you opened the document before you remembered opening it.

Nevertheless, there can be no doubt. You *must* have opened the document before remembering to have opened it. Why? Because of two things: memories are caused by the events remembered in them; and causes happen before their effects. Add the two together, and it follows that the opening must be temporally prior to the remembering. One of the premises of this argument is the claim that causes happen before their effects. Let us call this claim *C*. Now *C* is not an empirical fact. If the argument about the phenomenal content of experiences is right, we cannot know which experiences happen earlier and which happen later if we don't already assume the truth of *C*. So if *C* is known, as it seems to be, then *C* is known a priori.

But how can *C* be knowable a priori? One hypothesis might be that it is just a linguistic convention that whenever we have two events linked by the causal relation, we call the one that comes earlier 'the cause' and the one that comes later 'the effect'. But this hypothesis doesn't fit the story we have been telling so far: if it were true, we would not be able to use knowledge of causal order to make judgements about temporal order, since knowledge about temporal order would always be primary. So we have to explain the a priori status of *C* in a different way.

One option is that temporal relations reduce to causal relations (as causal theories of time claim). Another, perhaps more palatable option is that our temporal and our causal concepts are partly constituted through their interrelations, one of which is the identity of the temporal and causal order. In both cases, temporal concepts cannot be fully understood without reference to causal concepts and hence the presupposition of reduction would fail.

Obviously, this argument is quick and dirty and may be attacked from multiple directions. (See Gijsbers 2019 for a more elaborate defence.) But its aim here is not to convince, but to show that we cannot simply *assume* that temporal relations can be understood independently of our understanding of causation. The nature of the relation between time and causation is in fact one of the BPQs; as is the question whether Kant's arguments can undercut Hume's sceptical reasoning. So, once more, some of the BPQs already have to be answered before we can know whether reductive analyses of causation are even possible, hence, before we can even start using the Method of Examples.

The overall conclusion of these three sections, then, is this. The methodological superiority of the Method of Examples is supposed to lie in the way it solves the Meno problem, namely, by doing so in a non-circular fashion. We can judge theories of causation without first having to know which answers to the BPQs are correct. But theories that can be judged using the Method of Examples can only answer BPQs if they are reductive, and they can be only judged to be reductive given certain answers to the BPQs. Hence, there is a circle after all, and the methodological superiority of the Method of Examples is an illusion.

## 6 Intermezzo: related ideas

It is perhaps easy to mistake the approach of this paper for two related trends in the philosophy of causation: the anti-reductionism of Woodward (1990) and Carroll (1994, 2009); and the attack on the use of examples and counterexamples by Glymour et al. (2010). Before moving on to the conclusion, I wish to make the similarities and differences clear.

Woodward and Carroll both attack the idea that reductive analyses of causation are possible. To do this, they give examples that are supposed to show that causal facts do not supervene on non-causal facts. Carroll 1994 presents the example of the *perfectly rigid seesaw* (pp. 141–147). Hitting its left side from the top causes this side to go down, which in turn and simultaneously causes its right side to go up. Hitting its right side from the bottom causes this side to go up, which in turn and simultaneously causes its left side to go down. The movement of the seesaw is the same, but the causal order is different. Now Carroll suggests that in both scenarios, we consider another possible world that *starts* just after the hitting has occurred in the original world—it is so to say a ‘terminal segment’ of the original world. The two terminal segments, hence the two new worlds, can be completely identical in terms of what occurs; and yet, Carroll says, the causal order in them would be different. Hence, causal facts are primitive and do not supervene on non-causal facts.

If such arguments are successful, they would show that a reductive analysis is impossible. This conclusion is much *stronger* than what I have been arguing for. I have only been arguing that the judgement that a certain analysis is reductive requires a prior commitment to certain answers to the BPQs; I have not attempted to show that there is no coherent set of commitments under which reductive analyses are possible.

Indeed, my argument helps us to see one way in which the Carroll-Woodward strategy of giving counterexamples might fail. If the Method of Examples is circular in the way we described, then our commitments to certain answers to the BPQs may well play a role in determining which examples we accept as examples of genuine causation. Thus, the defender of a reductive analysis of causation might appeal to their commitment to reduction to argue that Carroll’s two worlds are identical and that there is therefore no difference in causal order. Of course, the strength of this appeal would depend on, among other things, the intuitive strength and clarity of the counterexamples. Either way, the claims of Carroll and Woodward are compatible with the theses of the current paper, but not implied by them.

It may appear that my approach is more closely tied to that of Glymour et al. (2010), who attack the idea that examples can be used as the standards for judging theories of causation. Their version of what I have called the Method of Examples, is what they call the Socratic Method, referring to Socrates’s attempt in the *Meno* to determine the nature of virtue by doing the following: “conjecture an analysis, seek intuitive counterexamples, reformulate the conjecture to cover the intuitive examples of the concept and to exclude the intuitive non-examples; repeat if necessary” (Glymour et al. 2010, p. 170).<sup>7</sup> Glymour and his co-authors argue that this inductive approach

<sup>7</sup> The fact that the *Meno* can be cited both as the source of the method of examples and counterexamples and as the source of criticism of that method perhaps serves to confirm Whitehead’s famous claim that all philosophy consists of footnotes to Plato.

to theories of causation cannot be successful, because the space of possible examples is far too big and far too varied: we could never go through a representative sample in order to really test our theories. In other words, they argue that even if some theory of causation gives the right answers to all of the standard examples found in the literature, this should give us very little confidence that it will also give the right answers in other, perhaps more complicated, cases.

Note that Glymour et al. do not claim that there is *no role* that examples and counterexamples can play. They merely wish to suggest that we cannot arrive at a theory of causation merely by testing proposed analyses against examples. The conclusions of the current paper fully agree with those of Glymour et al., but are based on very different reasons. Where Glymour et al. tell us that no set of examples can be *discriminating* enough to make the difference between good and bad theories of causation, I have argued that using examples to judge theories is only possible against the background of prior theoretical commitments about causation. Both arguments show that the Method of Examples cannot, by itself, suffice to tell us what the correct theory of causation is. For Glymour et al. this is primarily a *practical* problem—we cannot assess enough intuitive examples to make confident judgements—which therefore calls for a *pragmatic* solution—using Bayes nets to investigate real cases of actual causation. For me, there is a *fundamental* problem—the Method of Examples cannot solve the Meno problem—which calls on us to rethink how philosophical problems about causation can be answered. There is no conflict between our approaches, but the underlying arguments are very different.

## 7 Conclusions and prospects

I have argued, in Sect. 2, that there is a reason for the popularity of the Method of Examples: it promises to provide a non-circular criterion for judging theories of causation. This methodological presumption in favour of the Method of Examples led to the idea that a good theory of causation has to be, and has to be judged as, a reductive analysis of causal claims. But in Sects. 3 to 5 the promise turned out to be false; and there is no special reason to think of the Method of Examples as the ‘right’ way to judge theories of causation.

So what *is* the right way to judge theories of causation? I am extremely sympathetic to Phil Dowe’s suggestion, in the introductory chapter of Dowe 2000, that we can only judge such theories against the background of a particular philosophical project. Dowe shows at some length that theories are often criticised for not living up to commitments that their authors never undertook; e.g., theories that aim to establish a posteriori which physical processes in the world are really causal processes, end up being criticised for not giving a correct analysis of the everyday concept of causation. Rather, Dowe suggests, one ought to either criticise a theory for failing to live up to the demands of the project of which it is a part; *or* one ought to criticise that entire project.

Positively, this means that the success of a theory of causation depends on two things: its ability to play the requisite role within a particular philosophical project,

and the worthiness of the project itself. Taking the Method of Examples as our example, we can see that its applicability will depend widely on the context<sup>8</sup>:

1. Suppose that I am a Humean. Then I am committed to the claim that causation cannot be a fundamental ingredient of the world. But of course we talk about causation all the time. So what I want a theory of causation to do, is tell us how the kinds of causal statements that are important in daily life and in science can be reduced to the kinds of non-causal statements that fit my ontology. Hence, I embark on the kind of project that is familiar from the works of Lewis and his followers. Assessing examples and counterexamples will be an important element of my methodological toolkit, since I am looking for a reductive theory that captures the usual extension of the causal relation. So the Method of Examples is entirely appropriate for judging my theory.
2. Suppose that I am a philosopher in the mould of Clark Glymour or Peter Spirtes, conceiving of philosophy as a methodological discipline that creates tools for the sciences to use. What I really want a theory of causation to do, then, is to have practical success in identifying causal relations in real-life scientific data. Hence, I embark on a pragmatic project in which traditional examples and counterexamples play a very limited role—after all, *all* scientific diagnostic tools will have some false positives and some false negatives, so I don't consider counterexamples to my theory as necessarily fatal. So the Method of Examples is not very appropriate for judging my theory, although a large amount of counterexamples from within science *would* weigh against it.
3. Suppose that I am inspired by Kant to believe that time, space, causation and human thought are all inextricably linked. Then what I really want a theory of causation to do is to explicate the relations between these items. It would also be good if it allowed me to answer sceptical challenges from the Humeans. The Method of Examples will play no role in judging my theory, since it is very unlikely that my theory will have the form of an *analysis* of causation in terms of necessary and sufficient conditions. At least there is nothing in my philosophical framework that pushes me in that direction; as we can see from the fact that despite the crucial role of causation in Kant's philosophy, one will nowhere in his works find anything even remotely resembling a statement of the form "*c* is a cause of *e* if and only if..."

As Dowe asks us to do, then, the correct way to proceed in the philosophy of causation is to *always* keep an eye on the greater philosophical project within which a theory of causation is placed. This opens the doors towards a more pluralistic conception of the philosophy of causation. Here I do not mean the idea there is more than one type of causation (see Lombrozo 2010 for an example of such a claim), but that we may need different theories for different purposes. For instance, Dowe argues that we need to do both empirical and conceptual analysis. Our brief examples above suggest that,

<sup>8</sup> The three projects mentioned below do not exhaust the possibilities. For instance, one might embark on the more psychological project of giving an account of human causal reasoning. Here, close fit with the empirical facts and human cognitive capabilities would be very important; consistency and precision less so. Another approach might be to set out the ways in which causal beliefs and inferences function pragmatically in our scientific endeavours; an example of this would be Skyrms (1980).

for instance, a Humean reductive analysis (#1) and a more pragmatically useful theory to be used in the sciences (#2) could well exist next to each other in perfect harmony. But of course harmony is not universal. If the philosophical projects within which a theory of causation makes sense are directly opposed to each other, as the Humean (#1) and Kantian (#3) project would be, then one could hardly accept both theories of causation. But this is a battle to be fought not by assessing the theories themselves, but rather by assessing the larger frameworks of which they are part.

I believe that a more pluralistic attitude in the philosophy of causation will benefit approaches that have been unfairly subjected to the Method of Examples. Again, Dowe has made this case for his own physical causation project. In the remainder of this section, I wish to briefly look at two other examples: interventionism, and agency theories.

Take interventionism first. As I pointed out in Sect. 2, while it is an analysis of causation, it is not reductive. In fact, it contains a clear conceptual circle, and this has frequently been used as a reason to criticise the theory. Thus, Glymour (2004) and De Regt (2004) claim that such a non-reductive analysis cannot capture the meaning of causation. But this criticism makes sense only if Woodward is engaged in the Humean project (#1), or at least some other philosophical project that makes it important to have a reductive conceptual analysis. But Woodward is much more interested in understanding and improving the ways that we apply causal concepts in the sciences. Clearly, doctors engaged in finding out whether a proposed medicine causes recovery of patients are not working in a causal vacuum: they base their research on countless unproblematic causal claims, e.g., that ingestion of the pill causes the chemicals it contains to enter the patient's body. This also means that Baumgartner's (2009) criticism, namely that the interventionist analysis leaves it mysterious how causal inference can ever get started, misses the point as a criticism of Woodward's project. This is not to say that Baumgartner's point is uninteresting: it is very interesting to those who want to use interventionism to answer the big philosophical question of how humans are able to think in causal terms. But it gains traction only within this new philosophical context.

Now interventionism is relatively successful when assessed using the Method of Examples, so it doesn't have *that* much to gain from abandoning that Method as the gold standard within the philosophy of causation. Therefore, let us now turn our attention to a family of theories that fails miserably when assessed using the Method of Examples: agency theories of causation. These are theories that attempt to understand causation as dependent on agency, the most well-known examples being Collingwood 1940, Von Wright 1971 and Menzies and Price 1993.

Agency theories have never quite taken off, for reasons that are easy to understand when we recall the requirements generated by the Method of Examples. First, agency theories are extremely unpromising candidates for giving a *reductive* analysis of causation, since agency is itself surely a causal concept, as are related concepts like 'bringing about' or 'doing'. To this extent, such theories are in the same boat as interventionism. But added to this is the fact that agency theories are terrible at giving an extensionally correct analysis of causal claims. Imagine a theory vaguely like this:

$c$  causes  $e$  if and only if some agent  $A$  could bring about  $e$  by bringing about  $c$ .

Clearly, this analysis will fail as soon as we consider causes that lie beyond the power of any agent. Nuclear fusion in the sun causes plants on Earth to grow; but could any agent bring about the growth of plants by bringing it about that there is nuclear fusion in the sun? Clearly not, unless we explode the concept of ‘agent’ beyond all recognition. We can of course try to amend our theory, for instance by introducing the idea that causation also takes place in processes that are ‘analogous’ to agency (Menzies and Price 1993). But it is clear that an agency theory starts out with a massive disadvantage, since causation regularly happens where there is no agency in sight.

However, this disadvantage disappears as soon as we notice that the requirement to give an analysis of the form just mentioned is itself a requirement that depends on the philosophical context within which the theory is to be judged. Certainly, for the Humean who wishes to analyse away causation, giving such an analysis is the only form a theory of causation can take. But as we have seen, there is no general reason why a theory of causation should take that form. For instance, what is crucial about the theory of Menzies and Price is *not* some set of necessary and sufficient conditions for the truth of causal statements; rather, it is the idea that causation is like a secondary quality, and has reality only to beings that act in the world. Their theory should be judged on whether it manages to make this idea intelligible, whether it manages to integrate it in a broader philosophical framework, whether the theory generated in this way is fruitful in solving BPQs, and so on.<sup>9</sup> It is a pity that agency theorists have felt the need to present their theory in the form of an analysis, because that is what has made their approach seem hopeless. Freed from this requirement, it might be developed in a much more promising manner. (Something similar can be said for attempts to think causation in terms of powers, e.g., Molnar 2003 and Cartwright 2017.)

In sum, the Method of Examples is not a neutral, non-circular common ground which gives us an unproblematic set of criteria by which to judge theories of causation. Rather, we should recognise that theories of causation can be developed in many different philosophical contexts, and that these contexts come with their own criteria of success. By developing our sense of context, we can make the philosophy of causation both more diverse and more interesting.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

---

<sup>9</sup> A parallel might be useful here. Plato at one point presents the theory that knowledge is justified true belief; at another point, he suggests that all human knowledge is remembrance of an original knowledge that we forgot when we descended into this world. The former theory is an if-and-only-if analysis and can be criticised using Gettier-style examples. The latter is not an if-and-only-if analysis; and while it can certainly be criticised, it can’t be criticised using examples (unless one has independent access to the World of the Forms!). That is not to say that the second theory is independent of our actual applications of the concept of knowledge; rather, it presupposes our ability to apply the concept, without which it would have no content.

## References

- Allen, K. (2016). *A Naïve Realist Theory of Colour*. Oxford: Oxford University Press.
- Armstrong, D. (1983). *What Is a Law of Nature?*. Cambridge: Cambridge University Press.
- Baumgartner, M. (2008). Regularity theories reassessed. *Philosophia*, 36, 327–354.
- Baumgartner, M. (2009). Interdefining causation and intervention. *Dialectica*, 63, 175–194.
- Bird, A. (2007). *Nature's Metaphysics: Laws and Properties*. Oxford: Oxford University Press.
- Bird, A. (2016). Overpowering: How the powers ontology has overreached itself. *Mind*, 125, 341–383.
- Carnap, R. (1928). *Der logische Aufbau der Welt*. Leipzig: Felix Meiner Verlag.
- Carroll, J. W. (1994). *Laws of Nature*. Cambridge: Cambridge University Press.
- Carroll, J. (2009). Anti-Reductionism. In Helen Beebe, Christopher Hitchcock & Peter Menzies (eds.), *The Oxford Handbook of Causation*. Oxford University Press.
- Cartwright, N. (1979). Causal laws and effective strategies. *Noûs*, 13(4), 419–437.
- Cartwright, N. (2017). Causal powers: Why humeans can't even be instrumentalists. In J. D. Jacobs (Ed.), *Causal Powers*. Oxford: Oxford University Press.
- Collingwood, R. G. (1940). *An Essay on Metaphysics*. Oxford: Clarendon Press.
- Cooper, J. M., & Hutchinson, D. S. (1997). *Plato: Complete Works*. Indianapolis: Hackett Publishing Co.
- Davidson, D. (1967). Causal relations. *Journal of Philosophy*, 64, 691–703.
- Dowe, P. (2000). *Physical Causation*. Cambridge: Cambridge University Press.
- Earman, J. (1972). Notes on the causal theory of time. *Synthese*, 24, 74–86.
- Ellis, B. (2001). *Scientific Essentialism*. Cambridge: Cambridge University Press.
- Gijsbers, V. (2019). On the causal nature of time. In S. Kleinberg (Ed.), *Time and Causality across the Sciences*. Cambridge: Cambridge University Press.
- Gijsbers, V., & de Bruin, L. (2014). How agency can solve interventionism's problem of circularity. *Synthese*, 191, 1–17.
- Glymour, C. (2004). Critical notice of James Woodward, making things happen. *British Journal for the Philosophy of Science*, 55, 779–790.
- Glymour, C., Danks, D., Glymour, B., Eberhardt, F., Ramsey, J., & Scheines, R. (2010). Actual causation: A stone soup essay. *Synthese*, 175(2), 169–192.
- Grünbaum, A. (1968). *Modern Science and Zeno's Paradoxes*. Middletown: Wesleyan University Press.
- Heidegger, M. (1927). *Sein und Zeit*. Tübingen: Niemeyer.
- Hitchcock, C. (2001). The intransitivity of causation revealed in equations and graphs. *Journal of Philosophy*, 98, 273–299.
- Hitchcock, C. (2016). Probabilistic Causation. *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.)
- Hume, D. (1748). *An Enquiry Concerning Human Understanding*.
- Kant, I. (1781/187). *Kritik der reinen Vernunft*. Riga: Johann Friedrich Hartknoch.
- Kvart, I. (2004). Causation: Probabilistic and counterfactual analyses. *Collins, Hall, and Paul, 2004*, 359–387.
- Lewis, D. (1973). Causation. *Journal of Philosophy*, 70, 556–567.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell Publishers.
- Lewis, D. (2000). Causation as Influence. *Journal of Philosophy*, 97, 182–197.
- Lombrozo, Tania. (2010). Causal-explanatory pluralism: How intentions, functions, and mechanisms influence causal ascriptions. *Cognitive Psychology*, 61(4), 303–332.
- Mackie, J. L. (1974). *The Cement of the Universe: A Study of Causation*. Oxford: Clarendon Press.
- Mehlberg, H. (1980). *Time, Causality, and the Quantum Theory Vol. 1: Essay on the Causal Theory of Time*. Berlin: Springer.
- Menzies, P. (2014). Counterfactual Theories of Causation. In N. Z. Edward (Ed.), *The Stanford Encyclopedia of Philosophy*. Berlin: Springer.
- Menzies, P., & Price, H. (1993). Causation as a secondary quality. *British Journal for the Philosophy of Science*, 44, 187–203.
- Molnar, G. (2003). *Powers: A Study in Metaphysics*. Ed. S. Mumford. Oxford: Oxford University Press.
- Mumford, S. (2004). *Laws in Nature*. London: Routledge.
- Mumford, S., & Anjum, R. L. (2011). *Getting Causes from Powers*. Oxford: Oxford University Press.
- Oddie, Graham. (1982). Armstrong on the eleatic principle and abstract entities. *Philosophical Studies*, 41(2), 285–295.
- Paul, L. A., & Hall, N. (2013). *Causation: A User's Guide*. Oxford: Oxford University Press.

- Regt, H. de (2004). *Review of James Woodward, Making Things Happen*, Notre Dame Philosophical Reviews.
- Reichenbach, H. (1956). *The Direction of Time*. Berkeley: University of Los Angeles Press.
- Russell, B. (1913). On the notion of cause. *Proceedings of the Aristotelian Society*, 13, 1–26.
- Skyrms, B. (1980). *Causal Necessity*. New Haven: Yale University Press.
- Smart, J. J. C. (1969). Causal theories of time. *The Monist*, 53, 385–395.
- Suppes, P. (1970). *A Probabilistic Theory of Causality*. Amsterdam: North-Holland Publishing Company.
- Van Fraassen, B. C. (1970). *An Introduction to the Philosophy of Time and Space*. New York: Random House.
- Von Wright, G. H. (1971). *Explanation and Understanding*. Ithaca: Cornell University Press.
- Woodward, J. (1990). Supervenience and singular causal statements. *Royal Institute of Philosophy Supplement*, 27, 211–246.
- Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.