



CONTEMPORARY
DEBATES
IN PHILOSOPHY

CONTEMPORARY DEBATES IN
**PHILOSOPHY
OF SCIENCE**

EDITED BY CHRISTOPHER HITCHCOCK

Causes need not be Physically Connected to their Effects: The Case for Negative Causation

Jonathan Schaffer

Negative causation occurs when an *absence* serves as cause, effect, or causal intermediary. Negative causation is genuine causation, or so I shall argue. It involves no physical connection between cause and effect. Thus causes need not be physically connected to their effects.

10.1 Negative Causation

The terrorist presses the detonator button, and the bomb explodes. Here is *causation*.

But I have not said *which way* the detonator is wired. Perhaps pressing the button generates an electrical current that connects to the bomb and triggers the explosion:¹

Positive causation: pressing



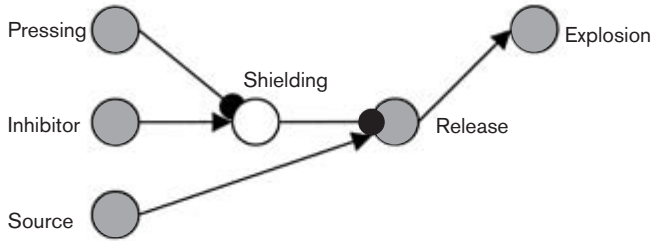
No pressing



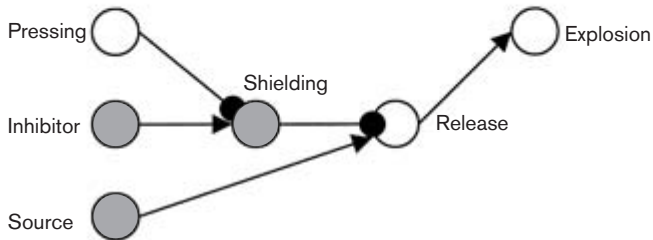
1 What follows are *neuron diagrams*, which are a useful way to represent causal structures. The conventions are as follows. Filled circles doubly represent firing neurons and occurring events. Unfilled circles doubly represent nonfiring neurons and absences. Arrows doubly represent stimulatory synapses and physical connections. Lines headed with black dots doubly represent inhibitory synapses and preventions (which are *not* physical connections in the relevant sense: section 10.2). If two neurons are connected by a stimulatory connection and the first fires, then the second will fire unless some other neuron inhibits it.

Or perhaps pressing the button causes the absence of an inhibiting shield that had been preventing the source current from triggering the explosion:

Negative causation: pressing



No pressing



If the detonator is wired as per the second diagram, then the case is one of negative causation, where an absence (the unfilled circle) serves as causal intermediary. Does it really matter, causally, which way the detonator is wired?

Whichever way the detonator is wired, the explosion is *counterfactually dependent* on the pressing. That is, the following counterfactual conditional is true: if the terrorist had not pressed the button, then the bomb would not have exploded. Such counterfactual dependence is sometimes thought to be constitutive of causation, and serves as the standard test for causation used in the legal system.

Whichever way the detonator is wired, the pressing is of a type that is *statistically relevant* to explosions in such circumstances. That is, the following probabilistic inequality holds: the probability of an explosion given the circumstances and a button pressing is greater than the probability of an explosion given the circumstances and no button pressing. Such statistical relevance is sometimes thought to be constitutive of causation, and serves as the basis by which statisticians infer causal relationships.

Whichever way the detonator is wired, the pressing serves as an *agential means* to achieve the explosion. That is, someone (for instance, a terrorist) whose end is to explode the bomb may manipulate the button to that end. Such agential means is sometimes thought to be constitutive of causation, and provides the original well-spring of our causal concept.

Whichever way the detonator is wired, the pressing may provide *predictive evidence* that the bomb will explode. That is, if one knows that the terrorist has pressed the button, then one is in position to know that the bomb will explode. (Likewise, the

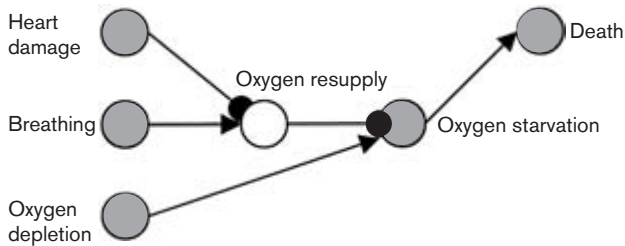
explosion may provide *retrodictive evidence* that the terrorist did press the button.) The scientist would be able to use the button pressing to predict an explosion. Such a prediction is a causal prediction, and thus would not be possible unless the pressing and explosion were causally related.

Whichever way the detonator is wired, the pressing may help to *explain why* the explosion occurred. That is, if one asks “Why did the bomb explode?” then someone may informatively answer “Because the terrorist pressed the button.” The government would certainly cite the button pressing to explain the explosion. Such an explanation is a causal explanation, and thus would not be possible unless the pressing caused the explosion.

Whichever way the detonator is wired, the terrorist is *morally responsible* for the explosion. That is, pressing the button exposes someone to blame (or praise) for the consequences of the explosion. The judge would hold the terrorist morally responsible, especially if anyone was hurt. Such responsibility is causal responsibility, and thus would not be possible unless the pressing did cause the explosion. In summary, whichever way the detonator is wired, all the central *conceptual connotations* of causation are satisfied, including counterfactual dependence, statistical relevance, agential means, inferential evidence, explanatory grounds, and moral responsibility.

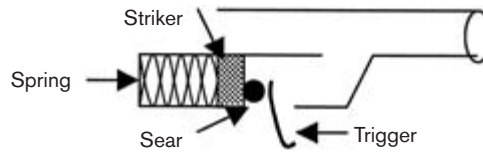
Still not convinced? The pattern of negative causation features in even the most *paradigmatically causal* cases. Suppose that the sniper feels murderous, pulls the trigger, fires a bullet through the victim’s heart, and the victim dies. Here is a paradigmatic causal sequence, *every step of which* is negative causation.

Workings backwards, *surely* the firing of the bullet through the victim’s heart causes the victim to die. But heart damage only causes death by negative causation: heart damage (c) causes an absence of oxygenated blood flow to the brain ($\sim d$), which causes the cells to starve (e). The mechanism of death can thus be represented as:

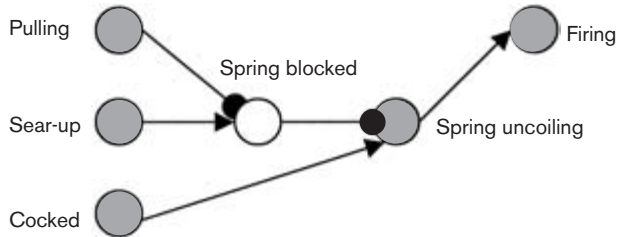


The Center for Disease Control (fully aware of the wiring) identifies heart disease as “the leading cause of death” in the United States.

At the next step backwards, *surely* the pulling of the trigger causes the bullet to fire. But trigger pullings only cause bullet firings by negative causation: pulling the trigger (c) causes the removal of the sear from the path of the spring ($\sim d$), which causes the spring to uncoil, thereby compressing the gunpowder and causing an explosion, which causes the bullet to fire (e). This may be seen in the following blueprint for a gun:

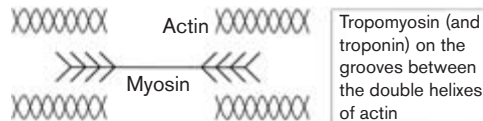


The mechanism by which bullets are fired can thus be represented as:

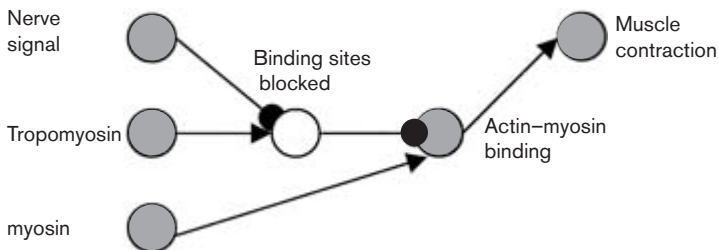


Even the National Rifle Association, which insists that “guns don’t kill people, people kill people,” concedes thereby that people who fire guns can cause death.

At the third and final step backwards, *surely* the sniper’s feeling murderous causes him to pull the trigger. But nerve signals only cause muscle contractions (such as that of the sniper’s trigger finger) by negative causation: the firing of the nerve (c) causes a calcium cascade through the muscle fiber, which causes calcium–troponin binding, which causes the removal of tropomyosin from the binding sites on the actin (*-d*), which causes myosin–actin binding, and thereby causes the actin to be pulled in and the muscle to contract (*e*). This may be seen in the following blueprint for a muscle fiber:



The mechanism by which muscles contract can thus be represented as:



Since all voluntary human actions are due to muscle contractions, it follows that *all voluntary human actions* (perhaps the most paradigmatic of all causes) involve negative causation.

Other paradigm cases of causation involving negative causation are cases of negligence and breach of contract, which are *explicitly recognized as causal by the law*. In this vein, H. L. A. Hart and Tony Honoré note:

There are frequent contexts when the *failure* to initiate or interrupt some physical process; the failure to provide reasons or draw attention to reasons which might influence the conduct of others; and the failure to provide others with opportunities for doing certain things or actively depriving them of such opportunities are thought of in causal terms. (1985, pp. 2–3)

Here is one of the many examples that Hart and Honoré then cite: “So a failure to deliver to a manufacturer on time a piece of machinery which he has ordered, may, like the destruction of the existing machinery, be held the cause of the loss of the profits which would have been made by its use” (1985, pp. 59–60).

Other paradigm cases of causation involving negative causation are cases that fall under the *causative verbs of ordinary language*. These include cases of removing, releasing, unveiling, unblocking, untying, unplugging, severing, disconnecting, letting go, cutting loose, switching off, and so on. In summary, negative causation features in the most paradigmatic cases of causation, as is supported by our intuitions, the law, and ordinary language.

Still not convinced? Negative causation is required by the most useful *theoretical applications* of causation. Saul Kripke (1972) proposes that *reference* is a causal notion: the reference of names is transmitted via causal chains. Kripke’s idea is that the reference of a name is not a matter of what description a speaker would associate with the name but, rather, of the causal chain by which the name was produced, transmitted, and ultimately entered into the speaker’s lexicon. It should be obvious that this transmission process is indifferent to positive versus negative causation. If a name is printed in a book, then its reference may be transmitted thereby, whichever way the printing press is wired.

Brian Skyrms (1980) and David Lewis (1981) argue that *rational decision-making* is a causal affair: one needs to calculate the expected effects of candidate actions. Skyrms’s and Lewis’s idea is that rational decision-making is not a matter of calculating the expectations *per se* but, rather, a matter of calculating the expected effects. It should be obvious that this calculation is indifferent to positive versus negative causation. If one’s overriding end is to explode the bomb, then one will act rationally by pressing the button, whichever way the detonator is wired.

Alvin Goldman (1977) maintains that *perception* is a causal affair: perception is a species of causal relation by which the world acts on the mind. Goldman explicitly notes that one can perceive black holes without any energy coming from them (pp. 281–2). In fact, black holes cause an absence of light that would otherwise be visible. Michael Tye then adds that the point extends to any perfectly black object: “This difficulty is not peculiar to astronomical contexts. It seems to me that perfectly black objects which are not too small or too distant may be seen with the naked eye provided that they are located against light backgrounds” (1982, p. 324). Thus the fact that one can see both white (positive) and black (negative) shows that perception is indifferent to positive versus negative causation. In summary, negative causation is

required by the most useful theoretical applications of causation, including applications to the theories of reference, decision, and perception.

Still not convinced? Negative causation is routinely recognized in *scientific practice*. Psychologists routinely invoke negative causation, such as in hormonal theories of sexuality:

The theory is that androgen causes masculine behavior and its absence causes feminine behavior . . . [M]ale rats were deprived of androgens by castration or by treatment with anti-androgenic drugs, which was seen to result in the later manifestation of the female pattern of lordosis . . . Thus a simple causal chain was established between the sexual behavior of animals such as guinea-pigs and rats and their hormonally sexually differentiated bodies. (Roberts, 2000, p. 2)

Biologists routinely invoke negative causation, such as in explaining disease. What causes scurvy is an absence of vitamin C, what causes rickets is an absence of vitamin D, what causes diabetes mellitus is an absence of insulin, and what causes dwarfism is an absence of growth hormone, and so on. The way in which HIV causes death is by disconnecting the immune system – an absence of functioning CD4+ T-cells serve as causal intermediary, allowing opportunistic infections and cancers to spread unchecked. One finds explicit discussion of the chains of negative causation involved, for instance, in Duchenne’s muscular dystrophy:

Muscular dystrophy in the *mdx* mouse has been described as a mutation in a colony of C57B1/10ScSn mice, which results in the absence of the 427 kDa membrane-associated protein dystrophin . . . A deletion on the human X-chromosome causes the absence of an analogous protein and leads to Duchenne muscular dystrophy (DMD) . . . The absence of dystrophin leads to the destabilization of [the transmembrane glycoprotein complex], yielding weaker muscle fibers that undergo progressive degeneration followed by massive necrosis. Ultimately, premature death of DMD patients occurs . . . (Caceres et al., 2000, p. 173)

Nor is the invocation of negative causation limited to psychology and biology. Chemists routinely invoke negative causation, such as in acid–base reactions. The addition of NH_3 (base) to H_2O (acid) causes the formation of OH^- (together with NH_4^+) because the H_2O loses a proton. Likewise, in oxidation–reduction reactions, reduction is understood as “a process in which the oxidation state of some element decreases,” as for instance when it loses O atoms. Physicists routinely invoke negative causation, such as in characterizing the process of “electron–hole pair generation”:

When an electron (which is a negative charge carrier) is freed from the atom, it leaves behind a *hole*, or the absence of an electron (which acts as a positive charge carrier). Free carriers are generated when electrons have gained enough energy to escape their bonds to the atom and move from the *valence band* to the *conduction band*. This process is called “electron–hole pair generation”. Electron–hole pairs can be created by any mechanism which delivers sufficient energy to an electron, including absorbing energy from light (as in a photo diode) and thermal excitation (absorbing heat energy). (Mason, 2000, p. 4)

Such electron–hole pair generation is routinely understood as causal: “The electron absence created by this process is called a hole.” And: “These positively-charged holes can cause a catastrophic negative shift in the threshold voltage of the device” (Wall and Macdonald, 1993). In summary, negative causation is routinely recognized in scientific practice, including throughout the practice of psychology, biology, chemistry, and physics.

Bringing the results of this section together, here is the full case for negative causation. First, negative causation is supported by all the central conceptual connotations of causation, including counterfactual, statistical, agential, evidential, explanatory, and moral connotations. Secondly, negative causation features in paradigm cases of causation including heart failure, gun firings, and all voluntary human actions, and is considered causal by the law and by ordinary language. Thirdly, negative causation is required by the most useful theoretical applications of causation, including applications to the theories of reference, decision, and perception. And, fourthly, negative causation is routinely recognized in scientific practice, including the practice of psychologists, biologists, chemists, and physicists. What stronger case can be made that *anything* is causal?

10.2 Physical Connection

David Hume (1748) glossed our naïve conception of causation as that of *necessary connection*. While Hume thought the connection not in the objects but, rather, projected by the mind, a number of subsequent philosophers have sought a causal connection in the objects via *physical connection* such as energy flow. These philosophers have maintained that causes need to be physically connected to their effects.

Negative causation refutes this program. Causation may well be objective, but it does not require physical connection.

There are three thematically related versions of the physical connections view. First, there is the idea that causation requires *transference*, of a property, or more specifically of energy-momentum. This idea has been developed by such philosophers as Jerrold Aronson (1971), David Fair (1979), and Hector-Neri Castaneda (1984).

Secondly, there is the idea that causation requires *processes*. This idea traces back to Bertrand Russell (1948). It was developed by Wesley Salmon (1984), who characterizes a causal process as a continuous qualitative persistence that is capable of transmitting a mark, of propagating structure. This idea was further developed by Phil Dowe (1992, 1995, 2000; see also Salmon, 1994, 1998), who characterizes a causal process as the world-line of an enduring conserved-quantity-bearing object.

Thirdly, there is the idea that causation requires an intrinsic *tie*. This idea has been developed by J. L. Mackie (1974). Douglas Ehring (1997) specifies this tie as the persistence line of a trope, and Max Kistler (1998, 2001) further develops this thought, while bringing it closer to Dowe’s view, by restricting the persisting tropes to those of conserved quantities.

These three approaches owe their distinctive aspects as much to historical pedigree as to philosophical difference. All understand physical connections as *lines of persistence*. They differ only in what is said to persist: unspecified for Russell and Mackie,

properties for Aronson, tropes for Ehring and Kistler, energy for Fair and Castaneda, structure for Salmon, and objects (those instantiating conserved quantities) for Dowe.

Thus the question of whether causes need to be physically connected to their effects becomes: Must there be lines of persistence from cause to effect? That is, must there be anything like energy flowing from cause to effect?

It is obvious that these research programs have no room for negative causation, since negative causation involves *no persistence line* between cause and effect. As Fair explains, “Omissions are non-occurrences. They obviously cannot be [physically connected] because, being only possible, non-occurrences cannot be the sources or the sinks of actual energy-momentum” (1979, p. 246). Thus when heart damage causes death, there will be no energy flow or other persistence line between the stopped heart and the starving brain. Rather, what is causally salient here is the *absence* of a physical connection. If only there had been a physical connection between heart and brain in the form of a flow of oxygenated blood, then the victim would have lived.

The physical connection view of causation may seem plausible if one concentrates on colliding billiard balls, or other cases of connection. But negative causation reveals the view to be a *hasty generalization*. Not all cases of causation involve physical connection. There is more than one way to wire a causal mechanism.

At this point, the road diverges. The physical connections theorist may *accept* the lesson of negative causation, grant that causes need not be physically connected to their effects, and hope that the notion of physical connection may still prove useful in some more complex account of causation. Fair takes this road, viewing the genuineness of negative causation as more or less beyond dispute:

I think a flagpole causes its shadow as an omission; the failure of the flagpole to transmit incident light causes the failure of the light to reach the shadow region. There is little question about the truth of this statement simply because it is very plausible that if the flagpole were removed, the light would reach the shadow region. Similarly, the ice on the road caused the auto accident because the road failed to transmit its usual frictional force to the tires. (1979, p. 248)

Fair then offers a more complex account of causation in terms of counterfactuals about physical connection. Causes, on this account, need no longer be physically connected to their effects. It is enough, to take the case in which both cause and effect are absences, that if the cause had occurred, then it *would have been* physically connected to an occurrence of the effect.

The other road open to the physical connections theorist is to *reject* the lesson of negative causation, insist that causes need to be physically connected to their effects, and swallow the consequences. Aronson takes this road, dismissing negative causation outright:

Consider a weight that is attached to a stretched spring. At a certain time, the catch that holds the spring taut is released, and the weight begins immediately to accelerate. One might be tempted to say that the release of the catch was the cause of the weight's acceleration. If so, then what did the release of the catch transfer to the weight? Nothing, of course. (1971, p. 425; see also Armstrong, 1999; Dowe, 2000, 2001; Kistler, 2001)

Such a response might have been tolerable were the consequences limited to the isolated case of launching a weight by a spring. But the consequences are not so limited. As shown in section 10.1, to dismiss negative causation is to swallow the following: counterfactual, statistical, agential, evidential, explanatory, and moral implications are not marks of causation; the folk are wrong that voluntary human action is causal, the law is wrong that negligence is causal, ordinary language is wrong that “remove,” “release,” “disconnect,” and so on are causal; philosophers are wrong that reference, decision, and perception are causal; and scientists are wrong that electron–hole pair generation and other negative processes are causal. I submit that no theory so dismissive deserves to be considered a theory of *causation*.

10.3 Meaning and Method

“When *I* use a word,’ Humpty Dumpty said in a rather scornful tone, ‘it means just what I choose it to mean – neither more nor less.’” If Humpty Dumpty were a physical connections theorist, he might have added “When *I* use the word ‘causes’ it means just physical connection – neither more nor less.” The sequel would of course be: “‘When I make a word do a lot of work like that,’ said Humpty Dumpty, ‘I always pay it extra.’”

So far, I have argued that negative causation proves that causes need not be physically connected to their effects. The argument that negative causation is genuine causation involved considering the conceptual connotations, paradigm cases, theoretical applications, and scientific practices concerning causation (section 10.1). But why should these considerations *matter*? That is, suppose that some philosopher simply *refused* to countenance negative causation, and insisted that causes must be physically connected to their effects, maintaining this thesis against even the most damning counterexamples, swallowing whatever absurdities might arise. “When I use the word ‘causes,’” this philosopher might say “it means just physical connection.”

Or suppose some truly crazed philosopher were to say “When I use the word ‘causes,’ it means just being over a mile apart.” What might one say to such a philosopher? One could point out that none of the connotations of causation – counterfactual, statistical, agential, and so on – require being over a mile apart. One could wheel out paradigm cases of causation in which cause and effect are not over a mile apart. One could point out that none of the theoretical applications of causation to reference, decision, and perception require being over a mile apart. One could mention that scientific practice in attributing causation does not require being over a mile apart. But if this philosopher remained *unmoved*, what more could one say?

Perhaps the crazed philosopher who defined “causes” in terms of being over a mile apart was merely engaged in *stipulation*. Such a stipulation may seem pointless and confusing, but it is still legitimate as such (especially so long as the word *is* paid extra). It is also of no further interest: mere *wordplay* cannot reveal the real structure of the world.

If this crazed philosopher were engaged in *description* of our actual causal concept, though, then he would have missed the mark by miles. There is an excellent method for explicating the meanings of our actual concepts. It is the method of *functional*

definition (Lewis, 1970). Roughly speaking, a functional definition of “causes” may be obtained by conjoining our most central platitudes involving the concept, replacing the term “causes” by the variable R , and uniquely existentially quantifying over the conjunction. The best satisfier of this definition then deserves to be considered the best candidate to be the meaning of our actual concept.

The central platitudes involving causation are to be drawn from the conceptual connotations, paradigm cases, and theoretical applications of the concept. Thus the functional definition of causation will look something like:

There exists a unique relation R such that: R is associated with counterfactual dependence & R is associated with statistical relevance & R is associated with agential means & R is necessary for inferential evidence & R is necessary for explanation & R is necessary for moral responsibility & . . . & R holds between heart damage and death & R holds between trigger pullings and gun firings & R holds between volitions and actions & . . . & R secures the reference of names & R is involved in rational decision & R is the genera of perception & . . .)

This is why the definition of “causes” in terms of being over a mile apart has no claim to descriptive adequacy. It violates virtually all the platitudes. And for *exactly* this reason, the definition of “causes” in terms of physical connection has no claim to descriptive adequacy, either. (Indeed, both definitions look to be about *equally* amiss.)

It may turn out, in the end, that no account of causation is perfect. Perhaps all violate some platitudes. Functional definitions allow for this possibility, by merely asking for *best* satisfiers. Negative causation is not merely a counterexample to the thesis that causes must be physically connected to their effects. There is a deeper point; namely, that any account of causation that requires physical connection between causes and effects is so far off the mark that it is *not even in the running* for the meaning of “causes.”

Not all philosophers will accept functional definitions as explicating the meaning of “causes,” though. The main alternative is to think of “causes” as a *natural kind term*, whose meaning (or essence or whatnot) is fixed, as a matter of a *posteriori* necessity, by the actual nature of causal relatedness. Many physical connections theorists hold this view, including Fair: “The hypothesized relationship between causation and energy-momentum flows is expected to have the logical status of an empirically discovered identity, namely that the causal relation is identical with a certain physically specifiable relation” (1979, p. 231; see also Aronson, 1982; Castaneda, 1984, p. 23). So just as chemists are said to have discovered that water is H_2O , so physicists are said to have discovered that causation is energy flow or some other species of physical connection.

To begin with, I believe that the natural kind term view of “causes” is just semantically wrong – “causes” is *not* a natural kind term in the way that “water” is. The mark of a natural kind term is the way in which we use it to describe counterfactual situations: given that the actual nature of natural kind k is x , we refuse to call counterfactual things that are k -like but not x by the k term. For instance, we refuse to use “water” to describe a hypothetical water-like liquid made of XYZ rather than H_2O (Putnam, 1975; see also Kripke, 1972). “Causes” does not work this way. We do not hesitate to

use “causes” to describe a hypothetical spell that works by magic rather than by energy. Indeed, the concept of causation belongs to the *nommic family* of concepts, alongside lawhood, explanation, disposition, and so on. *None* of the terms for nomic concepts have the characteristic counterfactual usage profile of natural kind terms.

But suppose for the sake of argument that “causes” is a natural kind term. And suppose that some philosopher insisted that the nature of causation is energy flow (or some other physical connection), dismissing even the most damning counterexamples as merely *intuitive*, ignoring whatever absurdities may arise as merely *conceptual*. “The nature of causation is energy flow,” this philosopher might intone “and no mere human intuitions or concepts can affect this.”

Or suppose that some truly crazed philosopher were to say “The nature of causation is being over a mile apart, and no mere human intuitions or concepts can affect this.” What might one say to such a philosopher? One could point out that none of the connotations of causation – counterfactual, statistical, agential, and so on – require being over a mile apart. One could wheel out paradigm cases of causation in which cause and effect are under a mile apart. One could point out that none of the theoretical applications of causation to reference, decision, and perception require being over a mile apart. One could mention that scientific practice in attributing causation does not require being over a mile apart. But if this philosopher remained *unmoved*, what more could one say?

Supposing, for the sake of argument, that causation is a natural kind, the question arises: *Which?* The crazed philosopher who identifies causation with being over a mile apart has obviously shot wide. Why so? There are three main views as to *which* natural kind a given term targets. The first view, Kripkean in spirit, is that a natural kind term refers to the dominant kind it was targeted at when the term was *born*. The Kripkean view can explain why the relation of being over a mile apart has no claim to be the actual nature of causation – the dominant kind that “causes” was originally targeted at includes events under a mile apart. And for *exactly* this reason, the Kripkean view rules that the relation of physical connection has no claim to be the actual nature of causation either. The dominant kind that “causes” was originally targeted at includes voluntary human actions without physical connection. Indeed, our term “cause” is the etymological descendant of the Latin “*causa*” and the Greek “*αἰτία*,” whose aboriginal use (which still survives in the law) is to identify grounds for moral complaint.

The second main view as to which natural kind a given term targets is the view, Putnamian in spirit, that a natural kind term refers to what the *community experts* now use it to cover. The Putnamian view can explain why the relation of being over a mile apart has no claim to be the actual nature of causation – the expert scientists in our community do not use “causes” in this way. And for *exactly* this reason, the Putnamian view rules that the relation of physical connection has no claim to be the actual nature of causation either. Psychologists, biologists, chemists, and physicists routinely use “causes” to cover negative causation, as when biologists speak of the absence of vitamin D as the cause of rickets, or when physicists speak of the causes and effects of electron holes in semiconductors.

The third main view as to which natural kind a given term targets is the view, Kaplanesque in spirit, that a natural kind term *rigidly* refers to the actual extension

of its functional definition. The Kaplanesque view can explain why the relation of being over a mile apart has no claim to be the actual nature of causation – the functional definition does not actually have this extension. And for *exactly* this reason, the Kaplanesque view rules that the relation of physical connection has no claim to be the actual nature of causation either. The actual extension of the functional definition of “causes” includes such cases of heart failures, gun firings, and voluntary actions, where there is no physical connection.

Of course, both the physical connections theorist and the crazed miles-away theorist might reject both functional definition and natural kind views. But the onus is on them, at this point, to say what possible conception of meaning could justify their theories. I conclude that there is no known and decent conception of meaning that could possibly *justify* the physical connections view (or even accord it *any* higher standing than the crazed miles-away view).

So far, I have spoken about meaning. But the situation would be essentially the same if one abjured the search for meaning, and simply proposed an *empirical specification* of actual-world causation in terms of actual physical connection, or in terms of actual distances of over a mile. The question would arise: *Why that?*

Dowe, for instance, has proposed to empirically specify actual-world causation as physical connection (understood in terms of the world-lines of objects that instantiate conserved quantities). But why *that*? Dowe is laudably explicit about his method. Empirical specification is to draw on science: “The empirical analyst can reply that there are procedures for investigating such an entity [as objective causation], namely, the methods of science, which is in the business of investigating language-independent objects. Empirical philosophy can draw on the results of science, . . .” (2000, p. 7). In particular, empirical specification is to draw on scientific judgments about the use of a term, so that it may be viewed as “a conceptual analysis of a concept inherent in scientific theories” (2000, p. 11).

It follows that Dowe’s method is similar to the Putnamian view mentioned above, in that it looks to the usage of the experts (but without the assumption that “causes” is a natural kind term). It follows for this reason that the relation of being over a mile apart has no claim to be the empirical specification of causation – the expert scientists in our community do not use “causes” in this way. And for *exactly* this reason, Dowe’s method rules that the relation of physical connection has no claim to be the empirical specification of causation either. Expert psychologists, biologists, chemists, and physicists routinely use “causes” to cover cases of negative causation.

To understand Dowe’s method better, it might help to understand why Dowe himself rejects alternative views such as statistical relevance. Here Dowe provides a *counterexample* involving a certain radioactive decay process (2000, pp. 33–40). Dowe’s counterexample is an extension of an *intuitive* counterexample involving a squirrel kicking a golf ball, but the radioactive decay version “is an idealization of a real physical nuclear decay scheme” (2000, p. 33), which the *physicist* Enge has glossed with the causal-sounding phrase “production process” (2000, p. 38). Well and good, I say. I only ask that the empirical philosopher not ignore intuitively paradigmatic cases of negative causation, and not ignore the testimony of psychologists, biologists when discussing the causes of disease, chemists, and physicists who speak of “electron-hole pair generation.”

10.4 Dowe's Arguments

In light of negative causation, why would anyone maintain that causes must be physically connected to their effects? That is, what possible considerations could convince someone that heart failures cannot cause death, that trigger pullings cannot cause bullet firings, that desires cannot cause actions, and so on? Likewise, what possible considerations could convince someone that decapitating someone (thereby stopping an influx of oxygenated blood) cannot cause that person to die?

Dowe (2000, 2001, and see chapter 9 of this volume) is perhaps the leading apologist for the physical connections view. Dowe deserves praise for being the first physical connections theorist to offer explicit *arguments* against negative causation, rather than merely biting so deadly a bullet. Dowe's overall strategy is to define a relation of "quasi-causation" in terms of merely possible physical connection, and then claim that negative causation is merely quasi-causal.

Dowe offers exactly two arguments for downgrading negative causation to the status of quasi-causation, the first of which is that he feels an *intuition of difference* between presence and absence cases:

There are cases where we have the intuition that an instance of quasi-causation is just that, "quasi", and not strictly speaking a genuine case of causation. "The father caused the accident by failing to guard the child." It's natural enough to use the word "cause" here, but when we consider the fact that the child's running onto the road was clearly the cause of the accident, but that the father did nothing to the child, in particular that the father did not cause the child to run onto the road, then one has the feeling that this is not a real, literal case of causation. So I claim that we do recognize, on reflection, that certain cases of prevention or omission are not genuine cases of causation. I call this the "intuition of difference". (2001, pp. 217–8)

Dowe recognizes that there are also cases, such as when chopping off someone's head *seems* to cause them to die, in which we have strong intuitions of genuine causation. But he criticizes the view that negative causation is genuinely causal as running afoul of the intuition of difference:

According to the genuinists – who say "quasi"-causation is genuine causation – the first intuition [the intuition of difference] can be dismissed but the second [the intuition of genuine causation] is important. Indeed, it would be fair to say that the main argument of the generalists (so far) is simply to cite intuitions of the second kind. However, without further explanation, this is an inadequate response to the existence of the intuition of difference, at least for those genuinists who purport to respect folk intuitions in their philosophical theorizing. Some account of the intuition of difference is required, and articulation and reiteration of the genuinist intuition does nothing to answer the point. (2001, p. 218)

In summary, Dowe maintains that our intuitions are conflicted here, and that only the quasi-causation view can explain this.

I offer three responses to Dowe's argument from the intuition of difference, the first of which is that Dowe has *misdescribed his own intuitions*. Look closely at the

first Dowe quote from the previous paragraph. Dowe begins with a case of parental negligence. He then immediately *acknowledges* that “It’s natural enough to use the word ‘cause’ here.”² But this is just to *acknowledge* that the case is intuitively *causal*. So much for intuitions.

What happens next in the passage under consideration is that Dowe presents a miniature *theoretical argument* to try to *overturn* our natural intuition of causality. This argument is not only beside the point as far as intuitions, but it is a cheat for two reasons. The first reason is that this argument implicitly assumes the physical connection view in sliding from “the father did nothing to the child” to “this is not a real, literal case of causation.” The corrective would be to note that by doing nothing the father thereby negligently *abandoned* the child, and that if he hadn’t the child would have been safe, which is what makes this a real case of causation. The second reason why Dowe’s argument is a cheat is that it turns on the misleading expression “the cause,” in “the child’s running onto the road was clearly the cause of the accident.” It is well known that this sort of locution invites people to make invidious distinctions between genuine causes.³ The corrective would be to acknowledge that the child’s running onto the road was *a cause* of the accident, as was the father’s negligence. Thus it comes as no surprise that Dowe reverses his natural intuitions at this point, driven by his theory and the invidious terminology of “the” cause. In summary, a close reading of Dowe’s own presentation reveals *no intuition of difference whatsoever*; rather, it reveals an intuition of genuine causality attacked by a theory-laden and misleading argument.

The second response that I offer to Dowe’s argument from the intuition of difference is that he has *understated the case for genuinism*, in two respects. First, genuinists do more than “simply to cite intuitions of the second kind.” The arguments of section 10.1 invoke not just paradigm case intuitions but also conceptual connotations, theoretical applications, and scientific practice. Moreover, as emerges in section 10.3, these are the considerations that determine the very meaning of “causes,” and that are supposed to guide Dowe’s own method of empirical specification. To be fair, Dowe explicitly limits his claim to “the genuinists (so far).” But in any case it should be clear that the genuinist can and should do far more than toss off a handful of intuitions.

The second respect in which Dowe’s argument from the intuition of difference understates the case for genuinism is that it treats all intuitions as if they were equal. But really intuitions come in *degrees*: some intuitions are paradigmatic near-certainties; others are mere borderline tugs. Forget the genuinist arguments from conceptual connotations, theoretical applications, and scientific practice, and consider only the genuinist argument from intuitions about paradigm cases, such as decapitation. Weigh this against whatever intuitions one might have about negligence, and

2 On this claim, Dowe and the law are in agreement. As Hart and Honoré note, “The use of this notion in the law is an extension of the general idea, common in non-legal thought, that the neglect of a precaution ordinarily taken against harm is the cause of that harm when it comes about” (1985, p. 195).

3 The misleading nature of “the cause” was first remarked on by J. S. Mill: “Nothing can better show the absence of any scientific ground for the distinction between the cause of a phenomenon and its conditions, than the capricious manner in which we select from among the conditions that which we choose to denominate the cause” (1846, p. 198).

here assume that one really does intuit that negligence is not genuinely causal (*contra* to what Dowe himself really intuits, and *contra* to what the law acknowledges). Would this be a case of “conflicting intuitions”? *Hardly*. This would be a case in which a paradigmatic near-certain judgment was being weighed against a tentative borderline tug. So even waving all the genuinist arguments but the paradigm case argument, and even granting Dowe’s alleged intuition of difference, this would at most provide a *negligible counterweight* to genuinism.

The third response that I offer to Dowe’s argument from the intuition of difference is that *genuinism does acknowledge a difference*. Positive and negative causation are different: the first involves physical connection and the second doesn’t. These are different ways of wiring a causal mechanism. All the genuinist denies is that this difference in the type of causal mechanism is a difference between genuine and fake *causation*. The physical connections theory has some initial plausibility (witness the fact that excellent philosophers such as Dowe defend it), although it turns out to be a hasty generalization. It is no surprise if we intuit that positive and negative causation are different: they are. It is no surprise that we are sometimes hesitant about whether this difference is a causal difference, because we are sometimes tempted to hasty generalization. (It is especially not a surprise if those who are most tempted to this hasty generalization are those who feel the most hesitancy.)

In summary, what Dowe has done in his first argument is to compare two cases of negative causation, one of which (decapitation) is a paradigm of genuine causation, and the other of which (parental negligence) is more borderline but still intuitively causal. Should one really conclude that it is *genuinism* that is in intuitive trouble here?

Dowe’s second argument for downgrading negative causation to the status of quasi-causation is that he sees negative causation as posing a *universal problem* for theories of causation:

So now we can outline a second argument against genuinism. At least for a number of significant theories of causation, allowing quasi-causation to count as causation leads to considerable theoretical difficulty. So there is considerable advantage to the claim that quasi-causation is not genuine causation – *viz.*, this solves those problems. (2001, p. 220)

Dowe then develops a detailed theory of quasi-causation in terms of counterfactuals concerning genuine causation, on which, for instance, the claim that the absence of *c* quasi-causes the absence of *e* is analyzed as: if *c* had occurred, then *c* would have caused *e*. Finally, Dowe uses his theory of quasi-causation to explain away genuinist intuitions. The explanation is that, because causation and quasi-causation are difficult to distinguish epistemically, and because they play the same roles conceptually, we get confused:

Further, our two results – the epistemic blur and the practical equivalence – furnish the answer to the objection that there are cases of quasi-causation which according to strong intuitions are definitely cases of causation, such as chopping off someone’s head causing death being a prevention since chopping off the head prevents processes which would have caused the person to continue living (the genuinist intuition). The answer is that the

epistemic problem and the practical equivalence together suggests that we take the intuition of similarity as an intuition that the case is either causation or quasi-causation rather than as an intuition that it is causation not quasi-causation. It's hard to see how folk could intuit the latter given the deep epistemic problem, and it's hard to see why folk would have such an intuition as the latter given the practical equivalence. (2001, p. 225)

In summary, Dowe maintains that quasi-causation solves theoretical difficulties while serving to explain away intuitions for genuinism.

I offer three responses to Dowe's argument from a universal problem, the first of which is that the problem simply is *not* universal. The only theories that Dowe cites as having a problem are Salmon's, Armstrong's, and Lewis's, the first two of which are just physical connections theories. Moreover, Dowe admits that "There are accounts of causation according to which cases of omission and prevention come out as clear cases of causation" (2001, p. 220). Here he cites the statistical relevance and agential manipulation theories. So the situation is that negative causation poses a problem for only some theories. What should one conclude from that? I would have thought that the logical conclusion would be that negative causation provides an argument *in favor of the theories that can handle it*. Should one really conclude, from the fact that negative causation is a counterexample to a few theories, that it is *negative causation* that is in trouble?

Moreover, I think a Lewis-style counterfactual theory has more resources to handle negative causation than has been recognized. Lewis considers three strategies, one of which (Lewis's third strategy: 1986, pp. 192–3) is to regard the absence description as a way of referring to a present event, so that "the father's negligence" is just a way of referring to his actual nap, or whatever he actually did. Lewis worries, though, that this does not fit his general counterfactual approach, because the effect might not depend counterfactually on whether or not the father napped. The worry is that, if the father hadn't napped, perhaps he would have merely read a book instead, and so merely have been negligent by another means. I think the solution is near at hand though, since we may take the use of the absence description to guide the interpretation of the supposition that the nap does not occur. In such a context, it is to be interpreted as the supposition of a watchful father, which delivers counterfactual dependence. If so, then the only theories that face difficulties here are Salmon's and Armstrong's, both of which are versions of the physical connections theory. If so, then Dowe's argument effectively reduces to the following: if negative causation is genuine causation, then it refutes the physical connections theory.

The second response that I offer is that Dowe's theory of quasi-causation is *unsuccessful*. It goes wrong in more complex cases in which negative causation is combined with *overdetermination*, in that there are multiple failures, each of which suffices for the effect. Suppose that the mason and the carpenter both fail to work on the building, as a result of which the building fails to be completed. Then the mason's failing to work ($\sim c1$) and the carpenter's failing to work ($\sim c2$) should each count as causes of the building's failing to be completed ($\sim e$).⁴ But neither $\sim c1$ nor $\sim c2$ even

4 This case is borrowed from Hart and Honoré:

[A] similar rule should apply as regards concurrent omissions, each sufficient to produce the harm. If a mason and a carpenter both fail to do their part of the work on a building, so that it cannot be completed

count as quasi-causes, on Dowe's treatment. For even if $c1$ had occurred, e would still not have occurred, due to $\sim c2$. (Likewise, even if $c2$ had occurred, e would still not have occurred, due to $\sim c1$.)

Dowe himself recognizes that his account of quasi-causation is troubled by a variety of complex cases, and in fact explicitly restricts his account to giving a *sufficient* condition. But what sort of solution fails to provide a necessary condition? And in what sense have genuinist intuitions been accommodated, by an account that only matches them in simple cases?

The third response that I offer is that Dowe's theory of quasi-causation is *semantically unstable*. Either Dowe's theory is the best solution to the allegedly universal problem of negative causation, or not. If not, then of course it proves nothing. But if so, then it deserves to be considered *the true theory of genuine causation*. For at that point Dowe will have offered us two relations, one (that of physical connection) which he labels "causation," and the other (that of counterfactuals involving the first) which he labels "quasi-causation." But don't be fooled by the labels: there is nothing "quasi" about the second relation.

For the sake of neutrality, I shall relabel Dowe's two relations "R1" (physical connection) and "R2" (counterfactuals about connection). Now consider: Which of R1 and R2 is more deserving of the label "causation"? This is a semantic question. There are methods for answering such questions (section 10.3). And R2 is the runaway winner by every measure. For it is R2 that squares with the central conceptual connotations of causation, R2 that covers all the paradigm cases of causation, R2 that underwrites the most useful theoretical applications of causation, and R2 that does justice to actual scientific practice.⁵

The "epistemic blur" and "practical equivalence" that Dowe trumpets are, if anything, even further evidence of genuineness. Suppose that Dowe's R2-theory is the best solution to the "universal problem" of negative causation. Then Dowe would have explicated one relation, R1, with none of the epistemic, practical, conceptual, intuitive, theoretical, or scientific markings of being the causal relation, and a second relation, R2, with all of the epistemic, practical, conceptual, intuitive, theoretical, and scientific markings of being the causal relation. In short, it would be R2 that paddles, waddles, and quacks like causation.

In summary, what Dowe has done in his second argument is to start with a problem that is not universal, partially re-solve it, and then stick the label "quasi" onto his solution. Should one really conclude that it is *genuinism* that is in theoretical difficulty here?

by the due date, and the building-owner loses an expected profit, each would be liable for the lost profit provided it was in their contemplation at the time of contracting. So, when a supplier of webbing delayed in delivering it to an overall manufacturer, and the latter incurred contractual penalties on the main contract, the jury was entitled to hold the webbing supplier liable to indemnify the overall manufacturer for the penalties paid, the failure to deliver the webbing being "a substantial factor" in causing the loss and "sufficient in itself to have delayed" the fulfillment of the main contract, despite the existence of other "contributory" causes of the default on the main contract. (1985, p. 236)

5 The attentive reader may recognize R2 as, in essence, Fair's modified theory of *genuine causation* (from the conclusion of section 10.2).

By way of conclusion, I submit that Dowe and other physical connections theorists are in a *dialectically impossible* situation. I mean, what would it take to convince you that decapitating someone cannot actually cause them to die? Let the physical connections theorist make what arguments they may. At the end of the day, their arguments will still need to be weighed against the fact that, by their theory, chopping off someone's head cannot cause them to die.

10.5 Misconnections and Differences

What is causation? While here is not the place to detail a positive theory, I would like to sketch one moral of negative causation. But first, it will help to have in mind a well-known counterexample to the *sufficiency* of physical connection for causation, namely the problem of *misconnection* (Hitchcock, 1995; Dowe, 2000; Schaffer, 2001). For instance, suppose that while Pam throws a rock at the window, Red sprays red paint in the air, so that the rock turns red on route to breaking the window. Then there is a physical connection from Red's hand to the rock over to the window, namely the line of red paint. But there is obviously no causation – Red's spraying red paint in the air did not cause the window to shatter. Or suppose that while Pam throws a rock at the window, Tim the innocent bystander gapes in horror. Then there is a physical connection from Tim's gaping to the window shattering, namely the lines traced by innumerable photons and other microparticles emitted from Tim and absorbed by the window. But there is obviously no causation – Tim's gaping did not cause the window to shatter.

So what is causation? What is it that positive and negative causation shares, and that misconnection lacks? The moral I would draw is that causation involves at least some aspect of *difference making*. In both positive and negative causations, whether or not the cause occurs *makes the difference* as to whether or not the effect will occur. For instance, the pulling of the trigger makes the difference as to whether or not the gun will fire, and the absence of blood flow makes the difference as to whether or not the victim will live. Misconnection is not causation, because it *makes no difference* as to whether or not the effect will occur. For instance, Red's spraying makes no difference as to whether or not the window will shatter, and Tim's gaping makes no difference there either.

Physical connections are often difference makers. *But not always*: some connections are of the wrong sort (Red's paint) or of the wrong magnitude (Tim's photon) to make a difference to whether the window shatters. *And not exclusively*: absences (victim's lack of blood flow) also work as difference makers. This is why the physical connections theory is plausible and often enough right, but no more.

The moral I would draw, then, is that causation has a *counterfactual* aspect, involving a comparative notion of difference making. I leave open whether causation is a purely counterfactual affair, or whether it involves some hybrid of counterfactuals and physical connections, as Fair suggests. Either way, negative causation proves that causes need not be physically connected to their effects.

Acknowledgments

I am grateful to Russ Colton, Phil Dowe, and Chris Hitchcock; and also to *Philosophy of Science* and the University of Chicago Press for allowing me to reprint material from my “Causation by disconnection” (Schaffer, 2000; © 2000 by the Philosophy of Science Association. All rights reserved).

Bibliography

- Armstrong, D. M. 1999: The open door: counterfactual versus singularist theories of causation. In H. Sankey (ed.), *Causation and Laws of Nature*. Dordrecht: Kluwer, 175–85.
- Aronson, J. L. 1971: On the grammar of “cause.” *Synthese*, 22, 414–30.
- 1982: Untangling ontology from epistemology in causation. *Erkenntnis*, 18, 293–305.
- Caceres, S., Cuellar, C., Casar, J. C., Garrido, J., Schaefer, L., Kresse, H., and Brandan, E. 2000: Synthesis of proteoglycans is augmented in dystrophic *Mdx* mouse skeletal muscle. *European Journal of Cell Biology*, 79, 173–81.
- Castaneda, H.-N. 1984: Causes, causity, and energy. In P. French, T. Uehling, Jr., and H. Wettstein (eds.), *Midwest Studies in Philosophy IX*. Minneapolis: University of Minnesota Press, 17–27.
- Dowe, P. 1992: Wesley Salmon’s process theory of causality and the conserved quantity theory. *Philosophy of Science*, 59, 195–216.
- 1995: Causality and conserved quantities: a reply to Salmon. *Philosophy of Science*, 62, 321–33.
- 2000: *Physical Causation*. Cambridge: Cambridge University Press.
- 2001: A counterfactual theory of prevention and “causation” by omission. *Australasian Journal of Philosophy*, 79, 216–26.
- Ehring, D. 1997: *Causation and Persistence*. Oxford: Oxford University Press.
- Fair, D. 1979: Causation and the flow of energy. *Erkenntnis*, 14, 219–50.
- Goldman, A. 1977: Perceptual objects. *Synthese*, 35, 257–84.
- Hart, H. L. A. and Honoré, T. 1985: *Causation in the Law*. Oxford: The Clarendon Press.
- Hitchcock, C. R. 1995: Salmon on explanatory relevance. *Philosophy of Science*, 62, 304–20.
- Hume, D. 1748: *An Enquiry Concerning Human Understanding*. New York: Doubleday, 1974.
- Kaplan, D. 1989: Demonstratives: an essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In J. Almog, J. Perry, and H. Wettstein (eds.), *Themes from Kaplan*. Oxford: Oxford University Press, 481–563.
- Kistler, M. 1998: Reducing causality to transmission. *Erkenntnis*, 48, 1–24.
- 2001: Causation as transference and responsibility. In W. Spohn, M. Ledwig, and F. Siebelt (eds.), *Current Issues in Causation*. Paderborn, Germany: Mentis, 115–33.
- Kripke, S. 1972: *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lewis, D. K. 1970: How to define theoretical terms. *Journal of Philosophy*, 67, 427–46.
- 1981: Causal decision theory. *Australasian Journal of Philosophy*, 59, 5–30.
- 1986: Causation. In *Philosophical Papers II*. Oxford: Oxford University Press, 159–213.
- Mackie, J. L. 1974: *The Cement of the Universe*. Oxford: Oxford University Press.
- Mason, A. 2000: Introduction to Solid State Physics. Found online October 12, 2002, at <http://www.engr.uky.edu/~ee562/562HO1-physics.PDF>

- Mill, J. S. 1846: *A System of Logic, Ratiocinative and Inductive; Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*. New York: Harper & Brothers.
- Putnam, H. 1975: The meaning of “meaning.” In K. Gunderson (ed.), *Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press, 131–93.
- Roberts, C. 2000: Biological behavior? Hormones, psychology, and sex. *NWSA Journal*, 12, 1–20.
- Russell, B. 1948: *Human Knowledge: Its Scope and Limits*. New York: Simon and Schuster.
- Salmon, W. 1984: *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.
- 1994: Causality without counterfactuals. *Philosophy of Science*, 61, 297–312.
- 1998: *Causality and Explanation*. Oxford: Oxford University Press.
- Schaffer, J. 2000: Causation by disconnection. *Philosophy of Science*, 67, 285–300.
- 2001: Causes as probability-raisers of processes. *Journal of Philosophy*, 98, 75–92.
- Skyrms, B. 1980: *Causal Necessity*. New Haven, CT: Yale University Press.
- Tye, M. 1982: A causal analysis of seeing. *Philosophy and Phenomenological Research*, 42, 311–25.
- Wall, J. and Macdonald, A. 1993: *The NASA ASIC Guide: Assuring ASICs for Space*. Pasadena, CA: California Institute of Technology Jet Propulsion Laboratory.

Further reading

- Anscombe, G. E. M. 1975: Causality and determination. In E. Sosa (ed.), *Causation and Conditionals*. Oxford: Oxford University Press, 63–81.
- Davidson, D. 1967: Causal relations. *Journal of Philosophy*, 64, 691–703.
- Hall, N. forthcoming: Two concepts of causation. In J. Collins, N. Hall, and L. A. Paul (eds.), *Causation and Counterfactuals*. Cambridge, MA: The MIT Press.
- Hitchcock, C. R. 2001: A tale of two effects. *Philosophical Review*, 110, 361–96.
- Lewis, D. K. 2000: Causation as influence. *Journal of Philosophy*, 97, 182–97.
- forthcoming: Void and object. In J. Collins, N. Hall, and L. A. Paul (eds.), *Causation and Counterfactuals*. Cambridge, MA: The MIT Press.
- Mackie, J. L. 1965: Causes and conditions. *American Philosophical Quarterly*, 2, 245–64.
- Paul, L. A. 2000: Aspect causation. *Journal of Philosophy*, 97, 235–56.
- Schaffer, J. 2003: The metaphysics of causation. *Stanford Encyclopedia of Philosophy*, online at: <http://plato.stanford.edu/entries/causation-metaphysics/>
- Tooley, M. 1987: *Causation: A Realist Approach*. Oxford: The Clarendon Press.