

Causal explanation

Bradford Skow

For the *Routledge Encyclopedia of Philosophy*

www.rep.routledge.com

Article summary

An explanation is an answer to a why-question, and so a causal explanation is an answer to “Why X?” that says something about the causes of X. For example, “Because it rained,” as an answer to “Why is the ground wet?,” is a causal explanation.

Causal explanation is philosophically important because explanation-in-general is philosophically important, and causal explanation is a basic kind of explanation. So a complete philosophy will include a theory of explanation, and a complete theory of explanation will offer criteria for being a causal explanation.

The simplest theory of causal explanation says that “E happened because X” is a causal explanation iff X describes one or more causes of E. Another theory permits explanations that do less to count as causal explanations; it is enough to say something about the causes of E, without identifying any particular cause. The “manipulationist” theory, by contrast, demands more. For any event E, there are factors which, had they been different, E would not have happened, or would have happened differently; this theory requires a causal explanation to both identify some of those factors and convey how E depends on them. No known theory is without problems.

If *causal explanation* is one basic kind of explanation, what are the others? Two candidates are *teleological explanation* (e.g. “the plant’s leaves turned east in order to face the sun”) and *reason-for-action explanation* (e.g., “that store sells milk; that was Smith’s reason for shopping there”). But it may be that explanations of these kinds are implicitly causal explanations—maybe, for example, every teleological explanation is equivalent to a causal explanation.

1. Why causal explanation is important.

An explanation, in the sense relevant here, is an answer to a why-question. For example, “The ground is wet because it rained” is an explanation, since it answers the question “why is the ground wet?” What, then, is a causal explanation? Philosophers disagree, but a rough starting point is that “P because Q” is a causal explanation only if Q says something about the causes of the phenomenon P describes. “The ground is wet because it rained” is a causal explanation, since the rain caused the ground to become wet.

Why is causal explanation philosophically important? One reason is that explanation-in-general is philosophically important, both because it appears in key philosophical questions (like, obviously, "what is the nature of explanation?"), and also because it is needed to answer other philosophical questions. For example, part of answering "What is the nature of science?" is articulating the aims of science, and it is plausible that aiming to discover explanations—aiming to figure out why things happen—lies in the nature of science.

It could be, however, that *explanation* is philosophically important, but the narrower category *causal explanation* is not. After all, *explanation first discovered in 1882* is narrower than *explanation*, but is certainly not philosophically important. What is the difference? *Causal explanation* is not just narrower than *explanation*; it is also a kind of explanation. The "narrower category/kind" distinction may be illustrated with a simpler category, *bird*. The category *bird first discovered in 1882* is narrower than *bird*, but it is a miscellany. The birds in it sharing no ornithologically interesting features and do not resemble each other any more than they resemble birds outside the category. By contrast *parrot*, while also narrower than *bird*, is not a miscellany; the parrot is a kind of bird. A scientific taxonomy of birds, one that groups together birds that belong together, will include *parrot* but not *bird discovered in 1882*. It is like that with *causal explanation*. Philosophers have been dividing explanations into kinds from the beginning. Aristotle is sometimes said to have distinguished four kinds of causes, but he is better taken to have distinguished four kinds of explanation, with causal explanation one of them. If *causal explanation* is one of a small number of basic kinds of explanation, then understanding it is part of understanding explanation in general.

2. Theories of causal explanation.

Three theories of causal explanation will be discussed here; they do not exhaust the field. I will assume that every causal explanation can be phrased as an answer to a question of the form "Why did E happen?", where E names or describes an event. Since on one legitimate use "event" names the kind of thing that can have causes, this is no real restriction.

First is the simple proposal that "E happened because Q" is a causal explanation if and only if Q identifies one or more causes of E. "The ground is wet because it rained" is a causal explanation on this theory.

Identifying a cause is certainly enough to make an explanation a causal explanation. David Lewis (1986) held that doing less than this can also be enough; he held that the first theory is too demanding.

Why might it be too demanding? You might answer with examples, but there are also theoretical reasons. In general, questions have both complete and partial answers. The complete answer to "Who came to the party?" will list all the people who came: Alice, Bob, Carol, and so on. But "Some people from Massachusetts came to the party" is still, in some sense, an answer to the question, even if not a complete answer; it is a partial answer. The distinction between complete and partial answers applies to answers to all kinds of questions, including answers to why-questions.

The simple theory fails as a theory of complete causal explanations. The complete answer to "Why did E happen?" must list all the causes of E; just listing one, or some, is not enough. Could the simple theory be right as a theory of partial causal explanations? It certainly identifies one way to be a partial answer. "The ground is wet because it rained" is not the complete answer to why the ground is wet, but it is true, and seems to owe its truth to identifying a cause of the wetness. The question is whether a partial answer needs to go so far as to identify a cause. If partial answers to why-questions pattern with partial answers to other questions, then they do not need to identify a cause. A partial answer to "who came to the party?" does not need to identify any particular attendee; similarly, a partial answer to "why did the party happen?" does not need to identify any particular cause of the party.

So what does it take to be a partial answer? David Lewis proposed that something is a partial answer to "Why did E happen?" if and only if it says something about the causes of E (Lewis 1986). Lewis intended this as a theory covering all explanations of events, but we may understand it as a theory of causal explanation. You can say something about E's causes without naming a particular cause, just as "some of them were from Massachusetts" says something about the people at the party without naming any particular attendee. Since as it stands the proposal is vague, Lewis provided a more precise definition of "says something about the causes."

Let us say that a *complete hypothesis about what the causes of E are* is a proposition of the form "C1, C2, ... were causes of E, and nothing else was a cause of E." If a bomb explodes, then "the lighting of the fuse was a cause of the explosion, and the burning of the fuse was a cause of the explosion, and the presence of oxygen in the room was a cause of the explosion, and nothing else was a cause of the explosion" is a complete hypothesis about the causes of the explosion (it is a false hypothesis, since the explosion had other causes, for example all the causes of the lighting of the fuse). Lewis suggested that Q says something about the causes of E if and only if Q rules out, or is incompatible with, at least one complete hypothesis about the causes of E.

On this theory, “the asteroid continued moving in a straight line because there were no massive bodies nearby” is a causal explanation, even if you hold that the absence of massive bodies did not cause the straight-line motion, because it rules out the hypothesis that the presence of a massive body was among the causes of the straight-line motion. Similarly, the theory says that “the gas’s temperature is 60 degrees C because any other temperature would violate the ideal gas law” is a causal explanation: if the ideal gas law had been false, the gas’s temperature would have had quite different causes (in general, whether one thing causes another depends on what the laws of nature are), and so a different complete answer to “why was the gas’s temperature 60 degrees C?” would have been true; this answer has been ruled out.

Lewis’s theory of causal explanation is more permissive than the theory that requires causal explanations to identify particular causes. Is it too permissive, counting as causal explanations things that are not explanations at all? If I break a window with a baseball bat, then “The window broke because there are no ghosts” seems false. But on Lewis’s theory it is (true and) a causal explanation: it rules out all complete hypotheses that say that ghostly activity was among the causes of the breaking.

One might defend Lewis by observing that “...because there are no ghosts” is uninformative—we all already know there are no ghosts—and so, since people usually ask “why?” to get new information, it is something that you should not assert. When we judge the statement false, the defense continues, we have confused falsity with unassertability.

It is, however, debatable whether we really are confused in this way. Also, not all problem cases can be solved by appealing to unformativeness and unassertability. Suppose quantum mechanics teaches us that radioactive decay is a spontaneous, uncaused process. Then Lewis’s theory says that “this radium atom decayed because its decay was uncaused” is a causal explanation, since it rules out all hypotheses on which the decay had causes. Yet it does not seem true; the fact that an event was uncaused seems to entail that it cannot be explained, rather than to itself be an explanation. Still, “its decay was uncaused” may be a surprising fact we did not know, so we are not here confusing falsity with unassertability.

If Lewis’s theory is less demanding than the simple theory, James Woodward’s manipulationist theory is more demanding. It holds that causal explanations “are explanations that furnish information that is potentially relevant to manipulation and control: they tell us how, if we were able to change the value of one or more variables, we could change the value of other variables” (Woodward 2003, 6). A causal explanation must answer at least some “counterfactual questions about the conditions under which [the phenomenon asked about] would have been different” (Woodward 2003, 191). On this theory “The ground is wet because it rained” is a causal

explanation because it teaches that the absence of rain is a condition under which the ground's wetness would have been different.

The manipulationist theory earns points for denying that "the window broke because there are no ghosts" is a causal explanation. Being told there are no ghosts provides no information about how the breaking could have been manipulated; it provides no answers to "how could the breaking have been prevented, or made to happen sooner, or more spectacularly?"

Is the manipulationist theory too demanding? In some (perhaps unusual) circumstances it appears to make simply identifying a cause not enough for a causal explanation. Suppose I break a window with a bat, and that unbeknownst to you my confederates were lying in wait to intervene and break the window when and how I did should I deviate in any way from the plan. "The window broke because I hit it with a bat" is surely (true and) a causal explanation, even if it is all you know about the breaking. But learning it leaves you ignorant of how the breaking could have been prevented, manipulated, or controlled. It does not answer any counterfactual questions about the conditions under which the breaking would have been different.

Of course learning this explanation might lead you to *think* that stopping me from swinging the bat, or enticing me to swing earlier, would have made a difference to whether, when, or how the window broke. But none of these would in fact have made a difference.

If you had tied up my confederates and taken away my bat, that would have made a difference: then the window would not have broken. But the explanation does not put you in a position to know this, or to know, of any action, that performing it was a way to manipulate the window-breaking.

3. The scope of causal explanation

If causal explanation is one basic kind of explanation, what are the others? Two candidates are *teleological explanation* and *reason-for-action explanation*. Teleological explanations cite ends, goals, or purposes, and are marked by the use of "in order to." In a biology class the instructor might gesture at a student, Smith, and ask "Why is Smith's heart beating?" One answer is certainly that Smith's heart is beating in order to circulate her blood. This answer describes one of the beating's effects, but not any of its causes. Thus teleological explanations appear to differ in kind from causal explanations.

To introduce reason-for-action explanations, suppose Jones and her friends are just beginning to learn number theory, and Jones, looking ahead to the class on prime numbers, has bet her friends

that there are infinitely many primes. Weeks later, after that class, Jones can be found celebrating, and you ask her why she is celebrating. Her answer: “Because there are infinitely many prime numbers!—that is my reason for celebrating.” It seems that what she says is true, but it does not describe any cause of her celebration. The fact that the primes are infinite cannot cause anything. Thus reason-for-action explanations appear to differ in kind from causal explanations.

However, it might be that teleological explanations and reason-for-action explanations are, after all, “implicitly” causal explanations, and so that *teleological explanation* and *reason-for-action explanation* are species of *causal explanation*. If so, then they are not basic kinds of explanation distinct from *causal explanation*.

Maybe, for example, it is necessary that “Smith’s heart is beating in order to circulate her blood” is true if only if (i) ancestors of Smith’s heart circulated blood, and (ii) this is a cause of Smith’s having a heart that circulates her blood. If so, then it is necessary that: Smith’s heart is beating in order to circulate her blood if and only if Smith’s heart is beating because her ancestors’ hearts circulated blood, where this “because” statement is a causal explanation. Then the teleological explanation is equivalent to a causal explanation.

Similarly, maybe it is necessary that “Jones is celebrating because there are infinitely many primes—that is her reason for celebrating” is true if only if (i) Jones believes that there are infinitely many primes, and (ii) this is a cause of her celebration. If so, then it is necessary that: Jones is celebrating because there are infinitely many primes—that is her reason for celebrating, if and only if Jones is celebrating because Jones believes that there are infinitely many primes, where this second “because” statement is a causal explanation. Then the reason-for-action explanation is equivalent to a causal explanation. These particular proposals about teleological and reason-for-action explanations are too simple, however, and it is controversial whether any attempt to show these kinds of explanation to be species of causal explanation, either building on these proposals or starting somewhere else, can succeed.

If teleological explanations and reason-for-action explanations are causal explanations, are any explanations non-causal? Certainly yes. Only an explanation of an event can be a causal explanation; and certainly “Why X?” often has an answer, even when X does not describe an event. At least some mathematical facts, for example, appear to have explanations, but those cannot be explanations of events. Moreover, it may be that some explanations of events are not causal explanations.

Classifying the varieties of non-causal explanation is an on-going enterprise. Two kinds of examples are worth mentioning. A certain pendulum has a period of 2 seconds; why? One

answer: because it is 100cm long, and in general the period T and length l of a simple pendulum are related by the law $T = 2\pi\sqrt{l/g}$ (g is the gravitational constant). Since the pendulum's length does not cause its period, this appears to be a non-causal explanation (Hempel 1965). It might be what some call a "metaphysical" or a "grounding" explanation: such explanations describe "more basic" facts in virtue of which the target fact obtains (Rosen 2010). Second example: if Smith has twenty-three strawberries and three children, and tries and fails to distribute the strawberries evenly among the children, then Smith failed because three does not divide twenty-three. This mathematical fact does not appear to say anything about the causes of the failure. Marc Lange says this is an example of a "constraint" explanation, which works by showing the phenomenon being explained to owe its occurrence to laws that exhibit a particularly strong variety of necessity (Lange 2016).

References and further reading.

Bennett, J. (1988) *Events and their Names*. Indianapolis: Hackett Publishing Company.
(A comprehensive discussion of theories of events.)

Hempel, C. (1965). *Aspects of Scientific Explanation*, New York: The Free Press.
(The most important post-positivist work on explanation.)

Lange, M. (2016) *Because Without Cause*. Oxford: Oxford University Press.
(Discusses explanation in mathematics, explanation by constraint, and other kinds of non-causal explanation.)

Lewis, D. (1986) "Causal Explanation." In *Philosophical Papers Volume II*, Oxford: Oxford University Press, 214-240.
(Defends a causal theory of explanation.)

Rosen, G. (2010) "Metaphysical Dependence: Grounding and Reduction," in B. Hale and A. Hoffmann (eds.) *Modality: Metaphysics, Logic, and Epistemology*, Oxford: Oxford University Press.
(Discusses grounding and explanation.)

Setiya, K. (2011) "Reasons and Causes." *European Journal of Philosophy* 19: 129-157.
(Argues that reasons-for-action explanations are causal explanations.)

Skow, B. (2016) *Reasons Why*, Oxford: Oxford University Press.

(Defends something like the simple theory, and argues that there is only one kind of non-causal explanation.)

Sober, E. (1983) "Equilibrium Explanation." *Philosophical Studies* 43: 201-210.

(Argues that equilibrium explanations are not causal explanations.)

Reutlinger, A. and Saatsi, J. (2018) *Explanation Beyond Causation*, Oxford: Oxford University Press.

(Recent collection of papers attempting to theorize non-causal explanation.)

Van Fraassen, B. (1980) *The Scientific Image*, Oxford: Oxford University Press.

(Argues that explanations are answers to why-questions.)

Woodward, J. (2003). *Making Things Happen*, Oxford: Oxford University Press.

(Defends a manipulationist theory of causal explanation.)

Wright, L. (1976). *Teleological Explanations*. Berkeley: University of California Press.

(Argues that teleological explanations are causal explanations.)