Levels of Reasons and Causal Explanation

Bradford Skow*

I defend the theory that the reasons why some event occurred are its causes. Many "counterexamples" to this theory turn on confusing two levels of reasons. We should distinguish the reasons why an event occurred ("first-level reasons") from the reasons why those reasons are reasons ("second-level reasons"). An example that treats a second-level reason as a first-level reason will look like a counterexample if that second-level reason is not a cause. But second-level reasons need not be first-level reasons.

1. A New Causal Theory of Explanation. It is obvious that some explanations of some phenomena speak of the causes of those phenomena. Simple examples come immediately to mind: the bridge collapsed because the wind reached a certain intensity, electrons flew off the metal because light shone on it. Much more controversial is the claim that every explanation of why some event happened must say something about the causes of that event. Hempel (1965, 352) proposed a counterexample to this claim, and philosophers have been proposing them ever since. But believers in the existence of noncausal explanations have to aim at a moving target, for different philosophers understand the claim that explanations must "say something" about causes in different ways. I have a new way to understand this claim, one that, I think, is the most natural way to understand it. I also think that the claim, understood my way, is true (with one qualification) and can be defended against the repeated objection that there exist noncausal explanations.

My theory starts with the idea, which has been held by many others, that explanations are answers to why questions.[1] A theory of explanation, then, should say what it takes for a proposition to be an answer to a why question.

*To contact the author, please write to: MIT Department of Linguistics and Philosophy, 77 Massachusetts Ave., 32-D808, Cambridge, MA 02139; e-mail: bskow@mit.edu.

1. Among those who hold that explanations are answers to why questions are Hempel (1965), with some qualifications; Van Fraassen (1980); and Bromberger (1992).

Philosophy of Science, 84 (December 2017) pp. 905–915. 0031-8248/2017/8405-0009$10.00
Copyright 2017 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/693966 Published online by Cambridge University Press

One standard form answers to why questions take is "P because Q": "The tide is high because the moon is overhead" answers "Why is the tide high?" But there is another form answers to why questions can take. The other form is "A/The reason why P is that Q."[2] Because answers and reasons-why answers are, in some sense, equivalent: "The tide is high because the moon is overhead" and "The reason why the tide is high is that moon is overhead" in some sense convey the same information. But I think that, for theoretical purposes, it is better to focus on reasons answers. (I argue for this claim in Skow [2016].)

A theory built around reasons-why answers will fill in the schema

1. A reason why P is that Q if and only if . . .

What should the claim that "explanations of events are causal" look like, if put into the form 1? Let "P" hold the place for a sentence that describes the occurrence of an event. Here is my proposal:

(T) A reason why P is that Q if and only if the fact that Q is a cause of the fact that P.

The same kinds of examples that lend credence to the idea that explanations of events are causal lend credence to its translation T into the language of reasons. The lighting of the fuse caused the bomb to go off; sure enough, it is also true that the reason why the bomb went off is that the fuse was lit. The electron's passing through a magnetic field caused it to accelerate; sure enough, the reason why it accelerated is that it passed through a magnetic field. But, the same examples philosophers have thought are counterexamples to the idea that explanations of events are causal also threaten to be counterexamples to T.

A bunch of these examples, I think, are based on the same mistake. There is a distinction to be made between "levels" of reasons. The examples fail because they confuse the two levels.[3] My aim in this article is to introduce the distinction and show how it can be used to defuse some examples. I will look, in particular, at Sober's (1983) claim that equilibrium explanations are noncausal and Lange's (2013) claim that "distinctively mathematical" explanations are noncausal.

2. I ignore here the forms used to give "teleological" explanations; I extend my theory to cover teleological explanations in Skow (2016).

3. I do think that one kind of counterexample succeeds against T: examples of "grounding" explanations. My true view is that every reason why a given event occurred is either a cause or a ground of its occurrence. But I ignore grounding explanation in this article.

https://doi.org/10.1086/693966 Published online by Cambridge University Press

2. Levels of Reasons. The distinction I want to introduce is that between

- a fact R being a reason why some event E occurred: then R is a "first-level" reason; and
- a fact F being a reason why R is a reason why E occurred: then F is a "second-level" reason, a reason why something else is a reason.

Reasons on the two different levels appear in answers to different why questions. The first-level reasons are the facts that belong in the complete answer to the question *why E occurred*. The second-level reasons, however, belong in the answer to a different why question: the question, concerning some reason R why E occurred, of *why R is a reason why E occurred*.

It is easy to come up with examples of first-level reasons. If I strike a match and, by striking it, cause it to light, then one reason why the match lit is that I struck it. What about an example of a second-level reason? We can find one by looking for the answer to the question of why the fact that I struck the match is a reason why the match lit. One answer is as follows: one reason why the fact that I struck the match is a reason why the match lit is that there was oxygen in the room at the time. (In general, background conditions to a cause's causing its effect are reasons why the cause is a reason why its effect happened.)

3. Second-Level Reasons Need Not Be First-Level Reasons. Here is the thesis about levels of reasons that I will defend:

> A fact can be a second-level reason without being a first-level reason. A fact F can be a reason why R is a reason why E happened, without F itself being a reason why E happened.

I say that F need not itself be a reason why E happened; I do not say that it cannot. The example I gave earlier shows that sometimes F is also a reason why E happened. The presence of oxygen, apart from being a reason why the striking of the match is a reason why the match lit, is also itself a reason why the match lit. But it is not always like this.

Here is an example in which a second-level reason is not also a first-level reason. Jill throws a rock at a window, Joan sticks out her mitt and catches the rock, and the window remains intact. The fact that Joan stuck out her mitt is a reason why the window remained intact. There is the first-level reason. Why is it a reason? The reason why it is a reason is that Jill threw a rock at the window. (If Jill had not thrown, certainly Joan's sticking out her mitt would not have been a reason why the window remained intact. The window would not have "needed" Joan's help.) But this second-level reason is not also a

https://doi.org/10.1086/693966 Published online by Cambridge University Press

first-level reason: that Jill threw a rock is not a reason why the window remained intact.[4]

In this case, the second-level reason that is not also a first-level reason is a fact that "corresponds" to the occurrence of an event: Jill's throwing of the rock. According to my theory T, first-level reasons why events occur all correspond to events, since they are all causes. But not all second-level reasons are like the two examples we have seen so far (Jill's throw, the presence of oxygen); not all second-level reasons correspond to events.

In fact, I hold that laws of nature are second-level reasons that are not also first-level reasons. If I drop a rock from 1 meter above the ground, and it hits the ground at a speed of 4.4 meters/second, the fact that I dropped it from 1 meter up is a reason why it hit the ground at 4.4 meters/second. The law relating impact speed s to drop height d, namely $s = \sqrt{2dg}$ (assuming drag is negligible and d is small), is a second-level reason: it is a reason why my dropping the rock from 1 meter up is a reason why the rock was going 4.4 meters/second when it landed. But it is not, in my view, also a first-level reason. It is not a reason why the rock hit the ground at 4.4 meters/second.

Mentioning laws of nature brings to mind Hempel's deductive nomological (DN) model of explanation, which says that an explanation of a fact F is a conjunction of facts that (i) entails F and (ii) essentially contains a law among its conjuncts (Hempel 1965). Hempel's theory is not framed as a theory of the reasons why facts obtain, but it is natural to interpret it as entailing that whenever there are any reasons why some fact obtains, at least one of the reasons is a law of nature. Hempel's theory is widely rejected, but I have a new diagnosis of where it goes wrong. Its mistake is to take certain second-level reasons, laws of nature, to also be first-level reasons.

I asserted without argument that laws are second-level reasons, but this is a natural view to have, on certain approaches to causation. One approach to causation takes laws to be central: whenever you have a cause and effect C and E, there are some laws connecting C to E—and C is a cause of E because of those connecting laws.[5] But that is just to say that whenever C is a cause of E, some law is a reason why C is a cause of E. Now I hold that when some fact F is a reason why C is a cause of E, then F is also a reason why C is a reason why E happened. So it follows from this theory of causation that laws are second-level reasons. If you start here, and in addition think that second-level reasons are always also first-level reasons, you head toward the char-

4. This is the kind of example many take to show that causation is not transitive; see Hitchcock (2001).

5. Hempel (1965, 349) endorses something like this idea about causation. It has had many other defenders.

https://doi.org/10.1086/693966 Published online by Cambridge University Press

acteristic thesis of the DN model, the thesis that among the reasons why some event happens is always at least one law. But this line of thought is fallacious because second-level reasons need not be first-level reasons, and, on my view, laws that are second-level reasons are never first-level reasons.

I admit that I have given no direct argument that laws are not first-level reasons. I would like to put the burden on the other side: why think they are? They are certainly second-level reasons: they are certainly reasons why causes are reasons why their effects happen. But as the Joan and Jill example shows, second-level reasons are not always first-level reasons. So why think they are in the case of laws? Certainly we have a sense that laws are "explaining something"; my view captures this sense, by assigning them the role of explaining why causes explain their effects. Why is that not enough?

4. How the Levels Can Get Confused. I said that the flaw in the DN model is that it misclassifies laws, which are second-level reasons, as first-level reasons. I also sketched an argument (with a false premise) that leads to this misclassification: "laws are second-level reasons, and second-level reasons are always first-level reasons, so laws are also first-level reasons." But I am not saying that Hempel or anyone else ever entertained this argument explicitly. Is there anything else to be said about how and why supporters of the DN model might have come to misclassify laws as first-level reasons?

Yes, there is. Pragmatic effects, effects of the rules of conversation on information exchange, can produce "data" that misleadingly suggest that laws are first-level reasons.

The reasons why an event happened are the parts of the answer to the question of why it happened. So if we come across a conversation in which one person asks "Why did E happen?" and another person answers this question by citing some fact F—and if that answer strikes us as correct—then we have some good evidence that F really is a reason why E happened.

Some of the evidence that laws are reasons why events happen appears to fit this pattern. Imagine someone walks into the room just as the rock hits the ground at 4.4 meters/second, and she sees that it hit at this speed (maybe the rock fell onto a device that measures impact speeds). A curious person, she asks me why it hit the ground at 4.4 meters/second. I respond, "Well, I dropped it from 1 meter up, and impact speed s is related to drop height d by the law $s = \sqrt{2dg}$ (and of course $\sqrt{2 \times 1 \times 9.8} \approx 4.4$)." Have I answered her question? And does the law that $s = \sqrt{2dg}$ appear in my answer? If so, then the law is a reason why the rock hit the ground at 4.4 meters/second—is it not?

If the answers to these questions are yes, yes, and yes, then, at least in some cases, a law is a reason why an event occurred. It is not hard to get from this conclusion to the claim (characteristic of the DN model) that this is so in

https://doi.org/10.1086/693966 Published online by Cambridge University Press

all cases and that when someone answers a why question without mentioning a law, her answer is incomplete.[6]

But the answers to these questions are not yes, yes, and yes. To explain what I think is going on I need to introduce another distinction: the distinction between a *good response* to a question and an *answer* to a question. If someone asks a question, obviously one good way to respond is to answer the question. But not every good response is an answer.

A simple example suffices to establish this. Sally asks whether Caleb is coming to the party. I know he is supposed to go to the party. I respond by saying, "He's sick." This is a good response. But it is not an answer. The only two possible answers are yes (he is coming) and no (he is not coming). I did not say either of those things.

There is a theoretical reason why we should expect there to be good responses that are not answers. The notion of an answer is a semantic one. But the notion of a good response is a pragmatic one. Whether a response to a question is good is a matter of what a cooperative speaker should say. In some circumstances, a cooperative speaker should respond to a question by doing something other than, or something more than, answering the question. In the simple example, I know that if I just answer the question by saying no, then Sally will immediately ask me why he is not coming. Since I can foresee that she will ask that, and since I know the answer to this question too, I respond to her explicit question not by answering it but by answering the expected follow-up question. It is okay in this case not to explicitly answer the question she asked because what I do say, my answer to the expected follow-up, conversationally implies that the answer to her explicit question is no.

I did not, however, need to be so indirect. I could have responded by answering both questions. I could have said, "no, he's sick." Here my response is good, but again it contains information that is not part of the answer to the question she explicitly asked. What keeps it from being a bad response is that the additional information is relevant to the topic of our conversation, and it is relevant because, although it is not an answer to her question, it is an answer to an expected follow-up question.

I think the same thing is going on in the dropped rock example. I responded to the question by saying "Well, I dropped it from 1 meter up, and impact speed s is related to drop height d by the law $s = \sqrt{2dg}$." My response is a good one, but it does not follow that every part of my response is part of an answer to the question asked. In my view, the first part of my response—"I dropped it from 1 meter"—is an answer to the explicit question ("why did the rock hit the ground at 4.4 meters/second?"), but the second

6. This "incompleteness" defense is most fully developed by Railton (1981). For one thorough argument against it, see Woodward (2003), chap. 4.

https://doi.org/10.1086/693966 Published online by Cambridge University Press

part, the law, is not. It, instead, is an answer to an unasked follow-up why question, a follow-up question I can anticipate would be asked immediately if I only answered the explicit question. The follow-up is, why is the fact that I dropped it from 1 meter up a reason why it hit the ground at 4.4 meters/second?

In summary: it is often a good thing to include a second-level reason in a response to the question why some event happened, but the fact that this is good thing to do is compatible with that second-level reason not being a reason why that event happened.

5. Equilibrium Explanations. I now have two distinctions: that between first- and second-level reasons and that between a good response to a why question and an answer to a why question. The two together provide the key to defusing many problem cases for T, the thesis that the reasons why something happened are its causes.

Sober argued that equilibrium explanations are not causal explanations. His main example of an equilibrium explanation was Fisher's answer to the question of why the ratio of males to females in the current adult human population is (very close to) 1 : 1 (Fisher 1931). "The main idea" of Fisher's answer, Sober reports, "is that if a population ever departs from equal numbers of males and females, there will be a reproductive advantage favoring parental pairs that overproduce the minority sex. A 1 : 1 ratio will be the resulting equilibrium point" (Sober 1983, 201). Parents who overproduce the minority sex are likely to have more grandchildren. So if males outnumber females in the population, the fitter trait is to be disposed to have more female children than male; being the fitter trait, this disposition should increase in frequency, with the result that the sex ratio is pushed from male-biased toward equality. The opposite happens if females outnumber males. Sober claims that this is not a causal explanation since "a causal explanation . . . would presumably describe some earlier state of the population and the evolutionary forces that moved the population to its present configuration. . . . Where causal explanation shows how the event to be explained was in fact produced, equilibrium explanation shows how the event would have occurred regardless of which of a variety of causal scenarios actually transpired" (202). In other words: Fisher's explanation does not say, for example, that the sex ratio in the year 1000 was such and such and that this caused the sex ratio in the year 1100 to be such and such, and so on. Instead it consists of a bunch of conditional facts: for each year in the sufficiently distant past, if the sex ratio in that year had had any "non-extreme" value (non-extreme meaning not all males or females), then the sex ratio today still would have been 1 : 1.

The first thing I want to say is that Sober makes a claim about what the causes of the current sex ratio are that I reject. He thinks that the only rele-

https://doi.org/10.1086/693966 Published online by Cambridge University Press

vant causes of the fact that the sex ratio is currently 1 : 1 are facts of the form *the sex ratio at time T is m* : *n*. I am with those who reject this claim. The fact that the sex ratio in 1000 was m : n is "too specific" to be a cause of the current sex ratio. There is a less specific fact, the fact that the percentage of males in 1000 was not 0% or 100%, that is as well placed to be the cause. The less specific fact is "better proportioned" to the effect than the more specific one, so it gets to be the cause.[7]

My disagreement with Sober might not seem to help much. Fisher's explanation is still a counterexample to T, is it not? Even if the cause of the current sex ratio is that the sex ratio in the past was never extreme, Fisher's explanation does not cite this cause either; his explanation instead contains a bunch of other facts, namely, the conditional facts described earlier. Does it follow that these conditional facts, which are not causes, are reasons why the sex ratio is 1 : 1 and thus that T is false?

I deny that those conditional facts that Fisher offers up are reasons why the sex ratio is 1 : 1. But I cannot just say this, for when Fisher offered those facts up in response to the question of why the sex ratio of 1 : 1, everyone celebrated his response; they did not reject it. How can his response be something to celebrate, if it did not answer the question?

The distinctions I introduced earlier show why. Fisher's response was something to celebrate because it was a good response to the question. But it can be a good response without containing an answer; in fact, that is exactly what I think is going on.

I think that the reason why the sex ratio is now 1 : 1 is that the sex ratio in the past was never extreme. But this is not something anyone would believe, or even be able to come to know, without an accompanying answer to the question of why that is the reason. So a good response to the question of why the sex ratio is now 1 : 1 must include an answer to the question of why the fact that the sex ratio was never extreme in the past is a reason why it is 1 : 1 now. And that is the question that the conditionals in Fisher's response constitute an answer to. Those conditional facts are second-level reasons why some other fact is a reason why the sex ratio is 1 : 1.

6. "Distinctively Mathematical" Explanations. Lange (2013) has recently described a class of explanations that he calls distinctively mathematical explanations and argued that they are not causal explanations. My interest is not in whether his examples qualify as noncausal by his criteria but in whether they are counterexamples to T. Here is one of his examples (also discussed in detail by Pincock [2007]):

7. A "proportionality requirement" on causation is defended in Yablo (1992) and Strevens (2008).

https://doi.org/10.1086/693966 Published online by Cambridge University Press

> Why did a given person [say, Jones] on a given occasion not succeed in crossing all of the bridges of Königsberg exactly once (while remaining always on land or on a bridge rather than in a boat, for instance, and while crossing any bridge completely once having begun to cross it)? . . . [Because] in the bridge arrangement, considered as a network, it is not the case that either every vertex or every vertex but two is touched by an even number of edges. Any successful bridge-crosser would have to enter a given vertex exactly as many times as she leaves it unless that vertex is the start or the end of her trip. So among the vertices, either none (if the trip starts and ends at the same vertex) or two could touch an odd number of edges. (Lange 2013, 488–89)

Here is what Lange says about why explanations like this one are not causal explanations: "these explanations explain not by describing the world's causal structure, but roughly by revealing that the explanandum is more necessary than ordinary causal laws are" (2013, 491). There is definitely something right, and deep, in what Lange says. But I do not think that his examples are counterexamples to T.

Let P be the property of bridge arrangements that a bridge arrangement has if and only if either every landmass or every landmass but two is met by an even number of bridges. The (supposed) answer to the question of why Jones failed that Lange presents boils down to this:

2. The bridges of Königsberg lacked P, and, necessarily, if a bridge arrangement lacks P, then no one can cross all the bridges exactly once (I am going to take Lange's qualifications about always remaining on land, etc., as given).

Now if 2 really is the answer to the question, then my theory is false. So is 2 the answer? There are two parts to 2. First is the fact that the bridges lacked P. Now it is no problem for my theory to recognize that this fact is a reason why Jones failed. For this fact is certainly a cause of his failure. The challenge to my theory comes if the second fact in 2 is a reason why Jones failed. For the second fact, that necessarily, no one can cross all the bridges exactly once, if the bridges lack P, cannot be a cause of Jones's failure.

I want to say the same thing about this example that I have said about the others. Statement 2, I maintain, is not an answer to the question of why Jones failed. It contains an answer as a part—the fact that the bridges lacked P. But it has another part, the necessary truth, that is not part of the answer. How is this compatible with the evident fact that 2 is a really good thing to say in response to the question of why Jones failed? Because the part of 2 that is not an answer to this question is an answer to an obvious follow-up why

https://doi.org/10.1086/693966 Published online by Cambridge University Press

question, namely, why is it that the bridges' lacking P is the reason why Jones failed?

Lange's diagnosis of this example, and the others he discusses, is quite sophisticated, and I do not have the space here to go in to all the things he says about them. Let me at least, however, mention one further thing he says. At one point he writes, "Even if [these examples] happen to appeal to causes, they do not appeal to them as causes. . . . Any connection they may invoke between a cause and the explanandum holds not by virtue of an ordinary contingent law of nature, but typically by mathematical necessity" (Lange 2013, 496). I am quite taken by this idea that an answer to a why question might appeal to causes but not appeal to them as causes. What might this mean, in terms of reasons why? Here is a natural suggestion: maybe in some cases a cause is a reason why its effect happened, but it is false that the reason why the cause is a reason why its effect happened is that it is a cause. The suggestion continues: cases like that are examples of "non-causal explanations."

I think the suggestion is plausible: if there truly are cases like that, they should be counterexamples to my theory. They are not, however, counterexamples to my theory as stated. I should amend my theory to make it more vulnerable:

> (T2) A reason why P is that Q if and only if (i) the fact that Q is a cause of the fact that P and (ii) the reason why the fact that Q is a reason why P is that the fact that Q is a cause of the fact that P.

Now the question is whether the Königsberg example, or any other example, is a counterexample to T2. I have a lot of thoughts about this but can only be brief here. Lange's idea is that since the "connection" between the bridges' lacking P and Jones's failure is secured by a mathematical truth (a theorem of graph theory), the bridges' lacking P, while a reason, is not a reason because it is a cause. I reject this claim. Even if the connection is secured by a mathematical truth, the cause is still a reason because it is a cause. This assertion requires defense, but I do not have the space here (I defend it in Skow [2016], chap. 5).

7. Conclusion. In this article I have presented a new causal theory of explanation that says that the reasons why an event occurred are its causes. I also drew two distinctions: that between the reasons why E happened and the reasons why those reasons are reasons, and that between an answer to a why question and a good response to a why question. I used these distinctions to defend the theory against the claim that equilibrium explanations and distinctively mathematical explanations are noncausal, and I believe the distinctions can be used to defend it against a wide variety of other examples.

https://doi.org/10.1086/693966 Published online by Cambridge University Press

REFERENCES

Bromberger, Sylvain. 1992. *On What We Know We Don't Know*. Chicago: University of Chicago Press.
Fisher, Ronald A. 1931. *The Genetical Theory of Natural Selection*. New York: Dover.
Hempel, Carl. 1965. "Aspects of Scientific Explanation." In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, 331–496. New York: Free Press.
Hitchcock, Christopher. 2001. "The Intransitivity of Causation Revealed in Equations and Graphs." *Journal of Philosophy* 98:273–99.
Lange, Marc. 2013. "What Makes a Scientific Explanation Distinctively Mathematical?" *British Journal for the Philosophy of Science* 64:485–511.
Pincock, Christopher. 2007. "A Role for Mathematics in the Physical Sciences." *Nous* 41:253–75.
Railton, Peter. 1981. "Probability, Explanation, and Information." *Synthese* 48:233–56.
Skow, Bradford. 2016. *Reasons Why*. Oxford: Oxford University Press.
Sober, Elliott. 1983. "Equilibrium Explanation." *Philosophical Studies* 43:201–10.
Strevens, Michael. 2008. *Depth*. Cambridge, MA: Harvard University Press.
Van Fraassen, Bas C. 1980. *The Scientific Image*. New York: Oxford University Press.
Woodward, James. 2003. *Making Things Happen*. New York: Oxford University Press.
Yablo, Stephen. 1992. "Mental Causation." *Philosophical Review* 101:245–80.

https://doi.org/10.1086/693966 Published online by Cambridge University Press