

Of Strawberries and Energy Conservation*

Bradford Skow

1. Marc draws a distinction between causal explanations and non-causal explanations, and argues that non-causal explanations exist. Marc has his own way of drawing the distinction; other philosophers draw the distinction differently. How does Marc draw it? He writes

The distinction between “causal” and “non-causal” explanations (as I will use these terms) lies in *how they work*—that is, in *what gives them explanatory power*. A “non-causal” explanation may incidentally identify (or, at least, supply information about) causes of what is being explained. But it does not derive its explanatory power by virtue of doing so. (3; my emphasis)

One variety of non-causal explanation is the “distinctively mathematical” explanation:

Distinctively mathematical explanations are “non-causal” because they do not *work by* supplying information about a given event’s causal history or, more broadly about the world’s network of causal relations. A distinctively mathematical explanation works instead (I will argue) roughly by showing how the fact to be explained could not have been otherwise. (5; my emphasis)

I find this talk of “how an explanation works” and “what gives an explanation explanatory power” a little obscure. But insofar as I do grasp it, I worry that some

*Forthcoming in *Metascience*, as part of a symposium on *Because Without Cause* by Marc Lange.

of Marc's paradigm examples do not actually satisfy his definitions. I'll say why I think this and suggest a way in which the definitions might be improved.

On page 17 Marc writes that "we might explain why visible light has a given speed in a given medium by the fact that electromagnetic waves within a certain ranges of frequencies have that speed there together with the fact that visible light consists of electromagnetic waves in that frequency range." He claims that this is a causal explanation: "it works by supplying relevant information about the world's network of causal relations. For instance, an explanation of light's speed in a given medium that appeals to the speed of electromagnetic waves there works by telling us that light's speed in that medium is caused by whatever causes the speed of electromagnetic waves there—and those factors cause light's speed by virtue of causing electromagnetic waves' speed." Here's a more concrete version of the example:

Q: Why does (visible) light travel at (roughly) 3×10^8 m/s in a vacuum?

A: Because visible light is a kind of electromagnetic wave, and all electromagnetic waves travel at 3×10^8 m/s in a vacuum.

Again, Marc says that the answer (A) "works by telling us" that light's speed is caused by whatever causes the speed of electromagnetic waves. Does it work by telling us that? It is hard to answer, partly because we're kind of on our own evaluating Marc's claim. There is no part of the book titled "How to figure out how an explanation works" (nor is there a part of the book that deserves that title). Marc does not propose any general theory of the form "'P because Q,' if true, works by" (In fact in an endnote that appears on page 404 he expresses skepticism about whether such a theory even exists.)

You might naively think that it is not hard at all to figure out how an explanation works: in all cases an explanation "P because Q" works by supplying the information that Q. Marc does not like this view: "[w]hether a given explanation gets its explanatory power from supplying information about the world's causal network or in some other way cannot be determined simply by identifying what information is included in the explanation" (404n16). *Why* doesn't Marc like it? Maybe just because many claims Marc finds intuitive about how this or that explanation works are incompatible with it. I can't find a more elaborate or principled

argument against the naive view anywhere in the book.

Anyway: (A) “works by telling us” that light’s speed is caused by whatever causes the speed of EM waves only if (A) *does in fact tell us* that light’s speed is caused by whatever causes the speed of electromagnetic waves. Does it? Well how could it? (A) is a proposition. Propositions don’t tell people things, in any ordinary sense of “tell.”

Set this issue aside; suppose (A) does tell us something about the causes of the speed of light. That’s not all it tells us: it also tells us that visible light is a kind of electromagnetic wave (etc). Does (A) really “work by” telling us the first thing, rather than by telling us only the second? Someone who had no concept of causation could still grasp the proposition expressed by (A), still know it to be true, and so still know why light travels at 3×10^8 m/s in a vacuum. Even if by learning (A) he has been “told” something about the causes of light’s speed, he didn’t “get the message”; he didn’t come to believe or know anything about those causes. Still, he knows why light travels at 3×10^8 m/s in a vacuum. Can an explanation really work by telling us some proposition if we can come to know that the explanation is correct without ever entertaining that proposition?

One simple example of an explanation that Marc thinks is a distinctively mathematical explanation is the “strawberries” example. Jane (not her real name) has some children and some strawberries, and tries to distribute the strawberries evenly among her children. She fails. Why? Marc accepts this answer:

That [Jane] has 3 children and 23 strawberries, and that 23 cannot be divided evenly by 3, explains why [Jane] failed when she tried a moment ago to distribute her strawberries evenly among her children without cutting any. (6)

Marc claims that

[t]his explanation works by supplying information about how [Jane]’s failure at her task, given nonconstraints understood to be constitutive of that task (as I discussed in chapter 1), comes to possess an especially strong variety of necessity. (139)

I think this claim about how the strawberry explanation works is false. It is worth emphasizing here that the explanation we are considering is not this one:

Jane failed to distribute her strawberries evenly among her children because doing so was mathematically impossible.

This explanation clearly “supplies information” about how Jane’s success was mathematically impossible (and so about how her failure is necessary in an especially strong sense), since the sentence following “because” *says directly* that success was mathematically impossible. But this explanation is *different from* the original explanation. “She failed because success was impossible” and “She failed because she had 3 children and 23 strawberries, and 3 does not divide 23” express different propositions. In the original explanation the sentence following “because” does not say directly that success was impossible; it merely says that Jane had 3 children and 23 strawberries, and that 3 does not divide 23. Since the original explanation does not supply the information that success was mathematically impossible by directly saying that it was impossible, how does it supply that information?

One might say that it supplies the information by entailing it: if I say that Jones is hungry, I don’t say directly that someone is hungry, but I do say something that entails this, and so in some sense I “supply the information” that someone is hungry. But the original explanation does not supply the information that success was impossible by entailing it either: the proposition that Jane had 3 children and 23 strawberries, and 3 does not divide 23, does not entail that dividing the strawberries evenly among the children (without cutting any strawberries) is mathematically impossible. The propositions that Jane had 3 children and 23 strawberries, and that *it is mathematically necessary* that 3 not divide 23 might entail that; but that is not what the original explanation says. (And Marc says in chapter 3 that explanations by constraint, while they need to include the constraint, need not in all cases include the fact that the constraint *is* a constraint.)

The kind of scenario that made me doubt Marc’s claim about the speed-of-light in a vacuum example makes me doubt Marc’s claim about how the strawberry example “works.” Here is Jones. Jones lacks all concept of modality. “Necessarily,” “possibly,” and similar words are just not in his vocabulary. He asks why Jane failed

and receives the original answer: because she had 3 children and 23 strawberries, and 3 does not divide 23. He believes what he hears, and comes to know it; he knows that she failed because she had 3 children and 23 strawberries, and 3 does not divide 23. He knows why she failed. Even if the original answer supplies the information that success was mathematically impossible, that information never arrived at its destination; Jones did not come to know, or believe, or even entertain the proposition that success was mathematically impossible. Still, as a result of hearing the explanation he knew why Jane failed. But if Jones can know why Jane failed without ever entertaining the information that success was mathematically impossible, how can it be that the explanation works by supplying that information?

In assuming that Jones lacks all concept of modality I am being overly-dramatic; it is not essential to the example. Maybe instead Jones is a crude empiricist: he believes that mathematical truths are knowable only on the basis of observational evidence, and he holds that all necessary truths are a priori. So he thinks that mathematical truths are contingent. He could still learn that Jane failed because she had 3 children and 23 strawberries and 3 does not divide 23, while denying that her failure was mathematically necessary.

2. Maybe I've misunderstood what Marc means by "works by." Let me suggest a way to capture at least part of what I think Marc had in mind by this phrase, using (what I think is) better terminology. Stop saying that whether an explanation is a causal one, or a distinctively mathematical one, is a matter of whether it "works by supplying information" of one kind of another. Instead say that whether an explanation is a causal one is a matter of *why it is an explanation*. My proposal is that Marc say that, while Jane's having 3 children was a cause of her failure (he accepts this), it is false that (Jane failed, in part, because she had 3 children) *because* her having 3 children was a cause of her failure (the parentheses are there to disambiguate). Instead, he should maintain that (Jane failed, in part, because she had 3 children) because her having 3 children helped make success mathematically impossible. In general, he could offer this definition: "P because Q" is a distinctively mathematical explanation iff (P because Q) because given Q, P is "strongly necessary," that is, has a variety of necessity that transcends the first-order laws of nature. (Similarly, "P

because Q” is a causal explanation iff (P because Q) because the proposition that Q is a proposition about some relevant causal facts.) Since one can know that P because Q without knowing why it is that P because Q, this definition makes room for the strawberries example to be a distinctively mathematical explanation even though Jones can know that Jane failed because she had 3 children (etc) without knowing that Jane’s failure was mathematically impossible. This definition makes room for saying that the facts that explain Jane’s failure are the facts that make her success impossible, while also saying that the fact that success *is* impossible is not (or need not be) among those facts.

3. Let’s work with this proposal. In light of it I want to discuss Marc’s claim that the strawberries example is a not a causal explanation. If the facts that explain her failure are (only) the facts that make success impossible, then surely the fact that 3 does not divide 23 is not among the facts that explain her failure. What makes success mathematically impossible is the fact that she has 3 children and 23 strawberries. The fact that 3 does not divide 23 doesn’t help make success mathematically impossible; it is, instead, the *reason why* another fact (the fact that she had 3 children and 23 strawberries) made success mathematically impossible. Right? But then the answer to the question why Jane failed is really “Jane failed because she had 3 children and 23 strawberries,” not “Jane failed because she had 3 children and 23 strawberries, and 3 does not divide 23”—though one might *say* the latter thing to abbreviate “Jane failed because she had 3 children and 23 strawberries, and *that is so because* 3 does not divide 23.”

If this is right, I wonder how Marc can rule out the hypothesis that this is a case of explanatory overdetermination. We’re contemplating the fact that Jane failed because she had 3 children and 23 strawberries; we’re wondering why this is so. Even if it is true that (she failed because she had 3 children and 23 strawberries) because 3 does not divide 23, that alone does not *preclude* its also being true that (she failed because she had 3 children and 23 strawberries) because her having 3 children and 23 strawberries caused her failure. Why not both? And if the second claim is right, then the strawberries example is a causal explanation after all (if both are right, it is an explanation of both kinds). If the second claim is wrong—well, by

what argument could it be shown to be false?

4. Maybe Marc's answer is that the fact that Jane had 3 children and 23 strawberries is "understood in the question's context as constitutive of the physical task . . . at issue" (33). What does this mean? Here is my interpretation: there is something Jane failed to do, and we have asked why she failed to do that thing; but if Jane had had 24 strawberries, and had divided *them* evenly among her 3 children, that would not have counted as doing the thing she failed to do that we asked about. This is maybe clearer if we explicitly distinguish between two things someone might do. One thing someone might do (this is an epistemic "might") is *distribute her strawberries evenly among her children*; another, different, thing is *distribute her 23 strawberries evenly among her 3 children*. These are different things that might be done, since one could do the first without doing the second. Since Jane failed to do either thing, the question "why did Jane fail?" is ambiguous. Are we asking why she failed to distribute her strawberries evenly (etc), or why she failed to distribute her 23 strawberries evenly (etc)? Since Marc takes her having 3 children and 23 strawberries to be "constitutive of the task," he must be addressing the second question.

Okay, so that's what I take the claim that the number of children and strawberries is "understood in the question's context as constitutive of the task" to mean; Marc proposes the *principle* that if a fact F is constitutive of the task asked about in the question of why E happened, then even if F both (i) is a cause of E and (ii) appears in an answer to the question why E happened, still it is false that that answer is an answer because F is a cause (F doesn't help explain why E happened "by virtue of" being a cause of E).

But now I don't see why this isn't just a bait and switch. Jane's having 23 strawberries is a cause of her failing to distribute her strawberries evenly, and appears in an answer to the question of why she failed; but her having 23 strawberries is *not* constitutive of the task of distributing one's strawberries evenly. So Marc's principle does not apply; he has not given us a reason to deny, in this case, that (she failed in part because she had 23 strawberries) because having 23 strawberries was a cause of her failure.

Now having 23 strawberries *is* constitutive of the task of distributing one's 23 strawberries evenly. His principle does apply to answers to the question of why she failed to do that. But for other reasons this cannot be the why-question he is focused on.

First notice that to the question of why Jane failed to divide her 23 strawberries evenly the answer is *not* "in part because she had 23 strawberries." Look at the whole "answer": "She failed to distribute her 23 strawberries evenly in part because she had 23 strawberries." This is no answer. (Make sure you give "her 23 strawberries" "narrow scope"; the claim we're evaluating is *not* "Jane's 23 strawberries are such that (she failed to distribute them evenly in part because there were 23 of them).") This is true but its truth is irrelevant; it is the answer to the question of why she failed to perform the task that have her having 23 strawberries is *not* constitutive of.) If the question presupposes that she has 23 strawberries, her having 23 strawberries cannot appear in the answer to the question.

What's more, Jane's having 23 strawberries is also not a *cause* of her failing to distribute her 23 strawberries evenly. To see why not consider a paradigm case of a "non-causal connection": the connection between Socrates' dying and Socrates' wife's becoming a widow. While the second depends on the first, this is not causal dependence. Why not?—after all, there is counterfactual dependence: if he hadn't died, she wouldn't have become a widow. The answer is that C cannot be a cause of E if the occurrence of a C-type event is a metaphysical precondition for the occurrence of an E-type event. A Socratic death is a metaphysical precondition for a widowing-of-a-Socratic-wife. Similarly, *having 23 strawberries* is a metaphysical precondition for *failing to distribute your 23 strawberries evenly among your 3 children*. (A Socratic death is not just a metaphysical precondition, it is metaphysically sufficient, while having 23 strawberries is not sufficient, just a precondition. This difference doesn't matter; the principle is about preconditions.)

Since Marc *does* take Jane's having 23 strawberries to be part of the answer to the why-question he is focusing on, and also a cause of her failure at the relevant task, he can't be focusing on the question of why she failed to distribute her 23 strawberries evenly. But, again, *this* is the only question that asks about a task that Jane's having 23 strawberries is constitutive of, and so the only question to which

Marc's principle is relevant.

5. New topic. In chapter 2 Marc claims that the explanation "every dynamical law conserves energy because it is a law that energy be conserved" is true iff the law of conservation of energy (is true and) constrains, places limits on, what the dynamical laws may be. Here is how he puts it:

The conservation law limits the kinds of forces there could have been to those that would conserve energy, and that is why every kind of force there actually is conserves energy. (50)

He elaborates:

the difference between energy conservation as a constraint and energy conservation as a coincidence is a difference in what is explanatorily prior to what. If energy conservation is a coincidence, then the various force laws are explanatorily prior to the law of energy conservation; they are partly responsible for its holding. On the other hand, if energy conservation constrains the force laws, then the conservation law is explanatorily prior to them. It does not suffice to determine the particular force laws there are. But it explains why every force law exhibits a certain feature. (50)

I like this distinction between energy conservation as a constraint, and energy conservation as a coincidence: in the latter case it is a law that energy is conserved merely because, as it happens, all the actual dynamical laws are conservative.

But *what is it* for the law of energy conservation to act as a constraint on the dynamical laws? Marc's answer is put in counterfactual terms: roughly speaking, the law of conservation of energy is a constraint iff, had the dynamical laws been different, conservation of energy would still have been true.

I think this answer is wrong.

Let's think about constraints. Joe's pub has a monitor standing outside who only permits people over the age of 21 into the room. In an intuitive sense, the monitor acts as a constraint on who is in the pub. If you ask why only people over

21 are in the pub, the right answer, or part of the right answer, is surely that there is a constraint in place on who may be in the pub, and that the constraint is that they must be over 21.

Now in this case, the existence of the constraint, of the monitor, entails the following counterfactual: if there had been other people in the pub, it still would have been true that all the people in the pub were over 21 (at least, this is so assuming that if anyone not in the pub had tried to enter, the monitor still would have enforced the constraint).

So the existence of the constraint entails the counterfactual; but it doesn't work the other way around. There is no monitor in front of the Harvard Bookstore. Anyone who walks by can go in. Suppose that lots of people do walk by, but just by coincidence the only people who go in are over 21. Imagine further that of the people who do *not* go in, only the people who are over 21 had any inclination to go in. The younger people who did not go in (again, this is just by coincidence) are bookstore-haters (some don't know how to read, others shop exclusively on amazon.com, . . .). It seems that in this scenario it is true that if there had been other people in the bookstore, it still would have been true that all the people in the store were over 21. But in this case there is no constraint on the age of the people in the bookstore operating.

Now it seems to me that the fact that all the dynamical laws conserve energy could be like the fact that everyone in the bookstore was over 21. It could be that it was invariant under the right counterfactual suppositions by coincidence, and so not because there is any constraint that "enforces" the requirement that all dynamical laws conserve energy. It could be that energy would still have been conserved, had the dynamical laws been different—but just by coincidence. In a scenario like that, the law of energy conservation does not in any sense act as a constraint on the dynamical laws. (Officially Marc's definition requires more counterfactuals to be true than the ones I just looked at; but I think the could all be true by coincidence as well.)

It seems to me that being a meta-law (in Marc's counterfactual sense) is not enough for a law to constrain the dynamical laws. What more is needed? Well in the case of Joe's pub it is not just true that had there been other people in the

pub, they would have been over 21; this is true *because there was a monitor at the door*. It seems that what more is needed, beyond the truth of the counterfactuals, is a certain kind of *reason why* they are true. So here is an alternative thesis:

For a law L to constrain the dynamical laws to have property P is for it to be the case that (i) every dynamical law has P because L is a law, and (ii) (if the dynamical laws had been different, it still would have been true that every dynamical law has P) because L is a law.

This seems right to me, but Marc will not want to accept it. For one thing, on Marc's view part of what it is for L to be a law is for the counterfactual in (ii) to be true; so his view is inconsistent with the claim that the counterfactual is true because L is a law. For another thing, this thesis spoils Marc's account of constraint explanation. Marc is proposing a philosophical account of what would be involved in the law of energy conservation explaining why all the dynamical laws conserve energy. I take a philosophical account like that to be a filling-in of the following:

- (Every dynamical law conserves energy because energy conservation is a law) because . . .

And the filling-in he proposes is

- (Every dynamical law conserves energy because energy conservation is a law) because the law of energy conservation constrains the dynamical laws.

But if my proposal about what it is to constrain the dynamical laws is true, then this becomes (assuming that "for it to be that X is for it to be that Y" and "P because X" entail "P because Y"):

- (Every dynamical law conserves energy because energy conservation is a law) because (every dynamical law conserves energy because energy conservation is a law, and [if there had been other dynamical laws, they would have conserved energy] because energy conservation is a law),

which is false, since a conjunction cannot explain one of its conjuncts.