THE JOURNAL OF PHILOSOPHY
VOLUME CXIII, NO. 2, FEBRUARY 2016

TIME ENOUGH FOR EXPLANATION*

There is an old Foster Brooks comedy routine in which Brooks appears on stage apparently drunk. He apologizes to the audience for standing before them in this condition but goes on to say that, in his defense, he has a good reason for being drunk: "I've been drinking all day."[1] This joke plays upon an ambiguity between levels of explanation that is of great philosophical interest. Brooks's audience has been set up to expect a high-level explanation of why he is drunk: some story of personal tragedy, perhaps. But what is delivered is a low-level explanation of the causal details of alcohol consumption. Both count as explanations, but in this case it is the high-level one that is expected and deemed to be informative.

Of course requests for explanations can be ambiguous and context sensitive in a variety of ways. But it is the distinction between high-level and low-level (or non-causal and causal, respectively) explanations that we wish to attend to in this paper. Familiar examples used to illustrate the distinction include the high-level (geometric) explanation of why a square peg will not fit into a round hole and the low-level (causal) explanation of why a particular attempt at inserting the peg fails.[2] A request for an explanation of the failure of the peg to slot into the hole is ambiguous between seeking the causal details and

*The authors would like to thank referees for comments on an earlier draft of this paper. The authors would also like to thank Alan Baker, Ben Blumson, Rachael Briggs, Otávio Bueno, Nikk Effingham, Aidan Lyon, Daniel Nolan, David Ripley, Nicholas J. J. Smith, and Alastair Wilson for useful discussion of this paper, as well as audiences at the Australasian Association of Philosophy Conference (2013), the New Zealand Association of Philosophy Conference (2013), the Sydney Time and Probability Workshop (2014), the Monash Workshop on the Metaphysics of Time (2012), the Swarthmore Mathematics and Philosophy Colloquium (2013), Victoria University Wellington (2014), and the University of Pennsylvania (2014).

[1] Foster Brooks at the opening of the Gerald R. Ford Presidential Museum in 1981.

[2] Frank Jackson and Philip Pettit, "Program Explanation: A General Perspective," *Analysis*, L, 2 (March 1990): 107–17.

0022-362X/16/1302/061–088

the geometry of the setup. More on this example and others like it shortly. For now we simply note that this ambiguity about levels of explanation is very much part of common sense (as the Foster Brooks joke highlights); it is not a controversial plank in some philosophical account of explanation. Moreover, as we shall see, both kinds of explanation play important roles in science.

The distinction between high-level and low-level explanations has been prominent in recent debates over mathematical explanation.[3] Here we find the debate focused on examples of mathematical explanations of physical facts (or *extra-mathematical explanations*, to be contrasted with *intra-mathematical explanations*, which involve the explanation of one mathematical fact by another). The mathematical explanations in question, like the square peg and round hole example, are high-level explanations. They abstract away from, or otherwise ignore, the causal details. For this reason, extra-mathematical explanation challenges popular, causal conceptions of explanation.[4] According to such conceptions, all (or at least most) explanation is

[3] Mathematical explanation has come to the attention of philosophers of mathematics recently for its role in arguments over the existence of mathematical entities. The indispensability argument makes a case for the existence of mathematical entities by appealing to the indispensable contribution such entities make to our best science. In order to defend this argument against various attacks, it has been argued by some that mathematics contributes to the explanatory power of current science (see Alan Baker, "Are There Genuine Mathematical Explanations of Physical Phenomena?," *Mind*, CXIV, 454 (April 2005): 223–38; and Mark Colyvan, "Mathematics and Aesthetic Considerations in Science," *Mind*, CXI, 441 (January 2002): 69–74). The explanatory turn in this debate has led to the identification of a number of putative cases of extra-mathematical explanation (see Baker, "Are There Genuine Mathematical Explanations," *op. cit.*; Mark Colyvan, *The Indispensability of Mathematics* (New York: Oxford University Press, 2001); Aidan Lyon and Mark Colyvan, "The Explanatory Power of Phase Spaces," *Philosophia Mathematica*, XVI, 2 (June 2008): 227–43; and Aidan Lyon, "Mathematical Explanations of Empirical Facts and Mathematical Realism," *Australasian Journal of Philosophy*, XC, 3 (September 2012): 559–78), and to a further debate on whether such cases provide substantial support for the existence of mathematical entities (see Alan Baker, "Mathematical Explanation in Science," *British Journal for the Philosophy of Science*, LX, 3 (September 2009): 611–33; Alan Baker and Mark Colyvan, "Indexing and Mathematical Explanation," *Philosophia Mathematica*, XIX, 3 (October 2011): 323–34; Sorin Bangu, "Inference to the Best Explanation and Mathematical Realism," *Synthese*, CLX, 1 (January 2008): 13–20; Joseph Melia, "Response to Colyvan," *Mind*, CXI, 441 (January 2002): 75–79; Joseph Melia, "Weaseling Away the Indispensability Argument," *Mind*, CIX, 435 (July 2000): 455–80; and Davide Rizza, "Magicicada, Mathematical Explanation and Mathematical Realism," *Erkenntnis*, LXXIV, 1 (January 2011): 101–14.

[4] The debate over mathematical and non-causal explanation, more generally, predates the recent interest. See, for example, Mark Steiner, "Mathematics, Explanation, and Scientific Knowledge," *Noûs*, XII, 1 (March 1978): 17–28; Mark Steiner, "Mathematical Explanation," *Philosophical Studies*, XXXIV, 2 (August 1978): 135–51; and Jack Smart, "Explanation – Opening Address," *Royal Institute of Philosophy Supplement: Explanation and Its Limits*, XXVII (March 1990): 1–19.

causal explanation.[5] Moreover, the question of whether or not mathematics can be genuinely explanatory has become important in the context of evolutionary biology. One recent shift in evolutionary biology is toward the identification and solution of local optimization problems, which can then be used to produce optimality models that both predict and explain the evolved traits of organisms. Indeed, according to Sutherland,[6] the next big step for biology generally is toward the development and implementation of optimality models of this kind. For the last two decades or so "optimality explanations" have been largely treated as causal explanations.[7] Recently, however, it has been suggested that optimality explanations are better thought of as non-causal explanations in which mathematics is doing most of the heavy lifting.[8]

There is, of course, some skepticism about the idea that mathematical facts can play a genuine role in the explanation of physical facts. This skepticism is important, and fuels ongoing debates, both within the philosophy of mathematics and within the philosophy of explanation more generally. It is not our aim, however, to defend a particular position on extra-mathematical explanation here. Instead, we focus on advancing an analogy between extra-mathematical explanations

[5] See, for instance, Nancy Cartwright, "From Causation to Explanation and Back," in Brian Leiter, ed., *The Future for Philosophy* (Oxford: Oxford University Press, 2004), pp. 230–43; Nancy Cartwright, *How the Laws of Physics Lie* (Oxford: Oxford University Press, 1983); and Bradford Skow, "Are There Non-Causal Explanations (of Particular Events)?," *The British Journal for the Philosophy of Science*, LXV, 3 (2014): 445–67. Woodward and Strevens might also be included in this list (see James Woodward, *Making Things Happen: A Theory of Causal Explanation* (Oxford: Oxford University Press, 2003); and Michael Strevens, *Depth: An Account of Scientific Explanation* (New York: Harvard University Press, 2008)). However, they remain officially neutral on whether there are non-causal explanations. Arguably, Lewis should also be included (see David Lewis, "Causal Explanation," in *Philosophical Papers, Volume II* (Oxford: Oxford University Press, 1986), pp. 214–40). Given Lewis's solution to the grandfather paradox, however, it is unclear whether he thought that there are non-causal explanations or not (see section 1.3 for discussion).

[6] See William J. Sutherland, "The Best Solution," *Nature*, CDXXXV (June 2005): 569.

[7] Proponents of this view include Potochnik, Orzack, and Sober. See Alison Potochnik, "Optimality Modeling and Explanatory Generality," *Philosophy of Science*, LXXIV, 5 (December 2007): 680–91; Steven Hecht Orzack and Elliot Sober, "How to Formulate and Test Adaptationism," *The American Naturalist*, CXLVIII, 1 (July 1996): 202–10; and Steven Hecht Orzack and Elliot Sober, "Optimality Models and the Test of Adaptationism," *The American Naturalist*, CXLIII, 3 (March 1994): 361–80.

[8] See Sam Baron, "Optimisation and Mathematical Explanation: Doing the Lèvy Walk," *Synthese*, CXCI, 3 (August 2014): 459–79; and Collin Rice, "Optimality Explanations: A Plea for an Alternative Approach," *Biology and Philosophy*, XXVII, 5 (May 2012): 685–703. Colyvan and Ginzburg and Ginzburg and Damuth argue that similar mathematical explanations can be found in ecology. See Mark Colyvan and Lev R. Ginzburg, "Analogical Thinking in Ecology: Looking Beyond the Disciplinary Boundaries," *The Quarterly Review of Biology*, LXXXV, 2 (June 2010): 171–82; and Lev R. Ginzburg and John Damuth (unpublished ms.), *Non-Adaptive Selection*.

and another kind of high-level explanation—what might be termed ‘extra-logical explanation’: the explanation of a physical fact by a logical fact. Specifically, a case of extra-logical explanation is identified that arises in the philosophical literature on time travel. This instance of extra-logical explanation is subsequently shown to be of a piece with extra-mathematical explanation. We argue that the analogy between the two kinds of explanation has implications for three areas: for the ongoing debate over extra-mathematical explanation, for the debate over time travel, and for the philosophy of explanation more generally. In particular, the analogy renders skepticism about extra-mathematical explanation harder to establish, sheds some light on a certain kind of solution to the grandfather paradox, and highlights the need for an account of scientific explanation that is capable of dealing with both causal and non-causal explanations.

I. EXTRA-LOGICAL EXPLANATION

Before considering cases of extra-mathematical and extra-logical explanations, it will be useful to first introduce a framework for thinking about higher- and lower-level explanation of the kind we are interested in. We should emphasize, however, that we introduce this framework largely for heuristic purposes, to facilitate our subsequent discussion of different kinds of explanation. We are not wedded to this way of thinking about the higher level/lower level distinction, or to any particular (metaphysical) way of understanding the framework in question. The framework we have in mind involves the distinction between process and program explanations.

I.1. Programs and Processes. A process explanation is an account of the actual causes that culminated in a particular explanandum. A program explanation, by contrast, is an explanation that appeals to some entity or property that is not itself causally efficacious but, rather, ensures the existence of whatever it is that causes the explanandum.[9] As Lyon notes, program explanations typically encode modal information that process explanations lack.[10] A program explanation can explain why a particular explanandum must be the case (for some appropriate modality), as opposed to why it is the case *de facto*. The latter is all that can be delivered by the process explanation.

An example will help. Consider the familiar round peg and square hole example.[11] Suppose that you are trying to push a square peg

[9] See Jackson and Pettit, “Program Explanation: A General Perspective,” *op. cit.*

[10] See Lyon, “Mathematical Explanations of Empirical Facts and Mathematical Realism,” *op. cit.*, pp. 9–11.

[11] See Hilary Putnam, “Philosophy and Our Mental Life,” in *Mind, Language and Reality: Philosophical Papers Volume II* (Cambridge, UK: Cambridge University Press, 1975), pp. 291–303; and Jackson and Pettit, “Program Explanation: A General Perspective,” *op. cit.*

through a round hole in a board, where the diameter of the hole is equal to the length of the side of the square peg. Clearly, you will fail, but why? There are two explanations. First, you fail to force the peg through the hole because, as a matter of fact, when you try to do so a particular part of the peg collides with a particular part of the board. Second, you fail to force the peg through the hole because of the squareness of the peg and the roundness of the hole (and their relative sizes); the squareness of the peg and the roundness of the hole ensure that each time you try to force the peg through the hole some part or other of the board will get in the way. You fail because what you are attempting is not possible.

The first explanation is a process explanation. It provides details of the particular causal processes that led to a particular failure. The attempt in question failed because *this* part of the peg bumped into *that* part of the board. Such an explanation lacks what might be thought of as modal robustness: it does not explain why you will always fail; it only explains why you failed this time. The second explanation is a program explanation. It appeals to certain properties that are not causally efficacious—namely, the geometric properties of the roundness of the peg and the squareness of the hole—in order to explain why you must fail. But note that this program explanation does not provide details of particular failed attempts, except to note that some part or other of the peg will always bump against some part or other of the board.

There is at present no general account of the nature of program explanation. Rather, the notion of a program explanation is typically elucidated by appealing to canonical cases, such as the peg/hole case outlined above. So perhaps some measure of caution is warranted.[12] Still, the distinction between process explanations and program explanations appears intuitive and offers a useful way to distinguish two quite different kinds of explanation. One way of elucidating what is important about program explanations proceeds via the notion of a structural constraint. Roughly speaking, structural constraints on a system are features that constrain the manner in which that system operates. The squareness of the peg and the roundness of the hole, for example, place important constraints on a closed system involving an agent, the peg, and the board. It is these structural constraints that

[12] Indeed, as Saatsi argues, it is difficult to make metaphysical sense of the program framework as applied to mathematical objects (see Juha Saatsi, "Mathematics and Program Explanations," *Australasian Journal of Philosophy*, XC, 3 (September 2012): 579–84). As noted, however, we are abstracting away from any particular metaphysical issues to do with the program/process distinction, using it primarily as a useful heuristic for pulling apart two kinds of explanation.

are facilitating the program explanation for why you cannot force the peg through the hole. These structural constraints (that is, constraints that ensure the existence of a given explanandum within a given system) will be important in what follows, and it is explanations that appeal to such constraints for which the phrase 'program explanation' will be reserved.[13]

It should be noted that the squareness of the peg and the roundness of the hole are not the only constraints relevant to the explanandum in the peg/hole case, although they are perhaps the most salient. That the hole is not elastic and that the peg is not pliable constitute further physical constraints on the system that contribute something to the explanans. That there is a cluster of such constraints is typical of explanations of this kind. This is also true for both the extra-mathematical and extra-logical cases considered below. But arguably all explanations are relative to such background assumptions, so there is nothing special about the explanations we have in mind in this regard.

When considering the distinction between a program and a process explanation, there is an ambiguity in the explanandum.[14] For instance, in the peg/hole case one might argue that there are really two explananda, one that is best explained by appealing to processes and one that is best explained by appealing to programming properties. These two explananda correspond to the following two "why" questions:

(A) Why did the peg, as a matter of fact, fail to go through the hole on a particular attempt?
(B) Why can't the peg ever go through the hole?

The process explanation appears to answer (A), while the program explanation answers (B), and so they seem to be distinct explanations corresponding to distinct explananda. Similar ambiguities arise for a number of explanatory cases that seem to admit of both process and program explanations. We always need to be clear about which "why" question we are seeking to answer.[15]

[13] We leave the notion of 'structural constraint' at an intuitive level for the moment. It will become clearer with the discussion of other examples where such constraints play a role. But the basic idea is that some outcomes are not possible (for some relevant modality), while others are rendered necessary (for some relevant modality). In the peg/hole example the structural constraints are (spatial) geometric, but the notion of structural constraint can be more general than this. One way to think of it is in terms of the structure of the phase space of the system—some outcomes are not accessible because of the geometric structure of the phase space. As we shall see, in at least some of the examples, this is precisely what the structural constraint in question amounts to.

[14] See Colyvan, *The Indispensability of Mathematics*, *op. cit.*, pp. 49–50.

[15] As in the Foster Brooks joke, mentioned in the introduction.

I.2. The Rings of Saturn and Balancing Knives. According to Lyon, (at least some) extra-mathematical explanations are instances of program explanations: such explanations appeal to certain non-causal features of the world—namely mathematical entities—in order to explain physical phenomena.[16] Let us consider a couple of examples.

The distinctive ring around the planet Saturn has a number of well-documented gaps. These gaps are disc-shaped regions where there are no (or at least, relatively few) orbiting particles. The existence of these gaps raises very natural questions about why such gaps exist and the locations and widths of these gaps.

The explanation of the gaps in the rings of Saturn is that they are determined by global properties of the system consisting of Saturn and its satellites.[17] An eigenanalysis of this system reveals that there are bands of unstable orbits, and the location and width of these unstable regions can be calculated. The mathematical details are not important for present purposes; the intuitive idea will do. Were a particle to be orbiting in one of the gaps, its orbital period would be such that it would pass near one of Saturn's satellites on a regular basis, and this would pull it out of the original orbit and into another. This process would repeat until the particle found itself in an orbit that did not suffer from close encounters with major satellites. In short, particles are pulled out of the unstable orbits and forced into stable ones. The unpopulated regions in the rings of Saturn are unstable in this sense: there is a kind of selection pressure against particles orbiting there.[18]

It is important to notice that the causal details are largely beside the point in describing and understanding the rings of Saturn; it is the mathematics of the eigenanalysis that does most of the explanatory work. Each particle in the system will have its own peculiar causal story of how it got to be where it is and why it is not in a gap region. But the eigenanalysis guarantees that there will be very few particles in the gap region and says why. What this analysis does, in effect, is reveal the structural constraints on the system, spelled out via the geometry of the phase space of the system. This ensures that stable orbits can exist only in certain regions. This is not to deny that there are causal stories and that in some contexts these stories might

[16] Lyon, "Mathematical Explanations of Empirical Facts and Mathematical Realism," *op. cit.*, pp. 9–10.

[17] See, for instance, Vladimir I. Arnol'd, *Huygens and Barrow, Newton and Hooke: Pioneers in Mathematical Analysis and Catastrophe Theory from Evolvents to Quasicrystals* (Basel: Berkhäuser, 1990).

[18] Colyvan and Ginzburg, "Analogical Thinking," *op. cit.*, call this process 'gravitational selection' and liken it to natural-selection explanations in evolutionary theory.

be important. It is just that, in the context of explaining the gaps in the rings of Saturn, these causal stories are at best only part of the story and at worst, misleading. They are misleading because if you were only in possession of the causal explanations of why each particle is where it is, the gaps in the rings would look like brute facts or accidents. But it is not an accident that the gaps are where they are and they have the width that they do. Given the setup of the system, the gaps had to be as they are. This important modal element of the explanation is provided by the eigenanalysis and is absent from the causal story.

There is nothing particularly exotic about this example; such stability explanations are commonplace in both science and everyday life. For example, the explanation of why you cannot balance a knife on its point proceeds along very similar lines. Suppose a knife is balanced on its point. It is sensitive to the smallest forces from any direction, so all it takes to upset the balance is for a particularly energetic or heavy gas molecule to hit the knife. The causal details would amount to a complicated account of the gas molecules in the region of the knife, the momenta of these particles, and so on. Such details may be important in explaining why the knife falls in the particular direction it does, but for the more general explanation of why you cannot balance a knife on its point, we need only the stability explanation. There will always be some causal story or other, but the details are of lesser importance.[19]

I.3. The Grandfather Paradox. We turn now to extra-logical explanation. Extra-logical explanations are controversial: it is not obvious that there are any such things and, moreover, it can be difficult to spell out exactly what they are, beyond what has already been said, namely that they are logical explanations of physical facts. To make a case for explanations of this kind, then, we proceed primarily by way of an example. Following the example, however, we will attempt to say a bit more about what extra-logical explanations might be. While we will not be offering any formal definition of 'logical explanation', we will sketch two potential accounts. Note that we are not wedded to either of these accounts: we offer them in a provisory mood only. Note also that something in the neighborhood of a theory of extra-logical explanation will be required to tell a completed story

[19] Alan Baker's example of prime life cycles of cicada in order to avoid predators can be couched in similar terms—in terms of stability selection—as can Lyon and Colyvan's (2008) example of galactic stability. See Baker, "Are There Genuine Mathematical Explanations," *op. cit.*, and Lyon and Colyvan, "The Explanatory Power of Phase Spaces," *op. cit.*

of explanation anyway. For it is clear that there are *intra-logical* explanations: explanations of logical facts by other logical facts. A theory of explanation, then, will be needed for explanations of this kind, and it is our view that, with such a theory in hand, it will be possible to generalize it to handle extra-logical explanations.

Our example of extra-logical explanation involves the grandfather paradox.[20] Suppose that Tim's grandfather died when he was very young, in 1985. In 2015 Tim builds a time machine and uses it to travel back in time to 1930, with the intention of killing his grandfather before his father was born. As Lewis argues, Tim cannot kill his grandfather in 1930, because his grandfather lived until 1985.[21] For Tim to kill his grandfather would be to change the past in an objectionable sense: it would be to bring it about that Tim's grandfather both lived and died in 1930. Worse still, if Tim were to succeed in traveling back in time and killing his own grandfather, then Tim's father would never be born and Tim would never be born to travel back in time to 1930. So if Tim kills his own grandfather, he brings it about that he both did and did not travel back to 1930, and he brings it about that he both does and does not exist.

Suppose that Tim tries anyway. What happens? He fails. Of course he fails. But why does Tim fail? One explanation is causal. Tim fails because of some commonplace reason: he is distracted by a noise, perhaps, or his gun jams, or he has a sudden change of heart.[22] This causal explanation, local to 1930, explains why Tim fails to shoot his grandfather *de facto*. There is a sense, however, in which this explanation is unsatisfactory, for it lacks modal force: although it explains why Tim fails *de facto*, it does not explain why Tim *must* fail.

If the local, causal explanation were the only explanation offered by Lewis, then it might well be argued that he failed to adequately resolve the grandfather paradox. It is possible to press the point by engaging in what might be thought of as "mystery mongering"

[20] David Lewis, "The Paradoxes of Time Travel," *American Philosophical Quarterly*, XIII, 2 (April 1976): 145–52, is the *locus classicus* for the grandfather paradox. For further discussion of the grandfather paradox see Larry Dwyer, "How to Affect, but Not Change, the Past," *Southern Journal of Philosophy*, XV, 3 (September 1977): 383–85; Larry Dwyer, "Time Travel and Changing the Past," *Philosophical Studies*, LXXVII, 5 (May 1975): 341–50; Paul Horwich, *Asymmetries in Time* (Cambridge, MA: MIT Press, 1987); Paul Horwich, "On Some Alleged Paradoxes of Time Travel," this JOURNAL, LXXII, 14 (Aug. 14, 1975): 432–44; Theodore Sider, "A New Grandfather Paradox?," *Philosophy and Phenomenological Research*, LVII, 1 (March 1997): 139–44; Theodore Sider, "Time Travel, Coincidences and Counterfactuals," *Philosophical Studies*, CX, 2 (August 2002): 115–38; Nicholas J. J. Smith, "Bananas Enough for Time Travel?," *British Journal for the Philosophy of Science*, XLVIII, 3 (September 1997): 363–89.

[21] Lewis, "The Paradoxes of Time Travel," *op. cit.*, p. 150.

[22] *Ibid.*

by supposing that Tim travels back in time repeatedly, trying to kill his grandfather each time, and constantly failing. Each failure is accompanied by a different causal explanation for why Tim fails. Stack all of these causal explanations up and the case looks bizarre: Tim is the subject of a great many gun jammings, changes of heart, distracting noises, and the like. Without a unifying account of why all of these events occur, time travel appears implausible.[23] This is precisely what worries some commentators on the grandfather paradox.

Fortunately, Lewis offers a second explanation for why Tim fails to kill his grandfather. Tim fails to kill his grandfather because if he succeeded, he would bring about a true contradiction, and contradictions are logically impossible.[24] What explains Tim's failure to kill his grandfather, then, is something about logic; specifically: Tim fails to kill his grandfather because the law of non-contradiction holds. This explanation has modal force. Because the law of non-contradiction is inviolable, Tim must fail to kill his grandfather. Moreover, this explanation is what unifies Tim's various and repeated failings. But this explanation is prima facie non-causal: it is not the case that the law of non-contradiction *causes* Tim to fail to kill his grandfather, as Lewis puts the point:

> The forces of logic will not stay his hand! No powerful chaperone stands by to defend the past from interference.[25]

Nevertheless, the logic is still explanatory and, moreover, explanatory in a manner that the causal antecedents of Tim's failing to kill his grandfather in 1930 are not.

The two explanations for why Tim fails to kill his grandfather line up very nicely with the program/process distinction. The first is a causal explanation that specifies what happens *de facto* to prevent Tim from killing his grandfather. This explanation is therefore a process explanation: it describes a causal process—such as the jamming of the gun, or a sudden change of heart—that prevents Tim from killing his grandfather. The second explanation for why Tim fails appeals to logical facts: the law of non-contradiction ensures that Tim will fail to kill his own grandfather. This is a program explanation: the law of non-contradiction places structural constraints on the entire

[23] Horwich, *Asymmetries in Time*, *op. cit.*, for example, takes this line. See Smith, "Bananas Enough for Time Travel?," *op. cit.*, for a response.

[24] See Lewis, "The Paradoxes of Time Travel," *op. cit.*, p. 150. For present purposes we set aside debates over the status of the law of non-contradiction. In any case, those who admit true contradictions are not at all tempted to admit contradictions such as those we are considering here.

[25] *Ibid.*, p. 149.

physical universe. The physical universe respects consistency, and that is why Tim fails to kill his grandfather and, indeed, why he must fail to do so.[26]

The grandfather paradox, we submit, is a case of extra-logical explanation: logical facts explaining physical facts. Moreover, we contend, the ongoing debate over the grandfather paradox either is due to a failure to fully appreciate the force of the extra-logical explanation Lewis gave us or else it is due to a rejection of the kind of explanation Lewis offered. Skepticism about Lewis's solution is the result of looking at the problem solely in terms of causal explanation. Doing so, however, ignores the impossibility of the task Tim is attempting. An important part of the explanation is being missed by tracking the causal histories, and this missing part is provided by logic. Logic does not cause the gun to jam, but it ensures that something or other will prevent the past from being changed.

We believe that, as with extra-mathematical explanations, extra-logical explanations are cases in which structural constraints are brought in to do high-level explanatory work—the structural constraint in this case being the law of non-contradiction. It could be argued, however, that the law of non-contradiction is not, in fact, a structural constraint. It is, rather, a *presupposition* for there to be any structural constraints at all. If the law of non-contradiction fails, it is tempting to think that pegs can fit into round holes after all; particles can stably orbit in the gaps in the rings of Saturn and knives can go on balancing on their points with ease. In short: give up the law of non-contradiction and all hell breaks loose; nothing constrains anything. If that is right, then the law of non-contradiction is just not up for grabs in the same way that, say, the structural constraints imposed by geometry in the peg/hole case might be.

But the law of non-contradiction *is* up for grabs. Indeed, as Priest has argued it may in fact be false (at least on some ways of formulating it).[27] Fortunately, there are paraconsistent logics available—Priest's *LP* is one—in which contradictions can be entertained without trivialization. The viability of these non-explosive logics shows that it is possible to give up the law of non-contradiction without being forced

[26] It is worth noting, however, that the law of non-contradiction is not thereby rendered trivial. It is a mistake to think that because the law of non-contradiction is such a wide-ranging constraint, it is no constraint at all. Here it structures the outcomes into those that are consistent and those that are not, with the latter ruled out. There are non-trivial mathematical and scientific theories that do not respect consistency (for example, naïve set theory and Bohr's theory of the atom).

[27] See Graham Priest, *Beyond the Limits of Thought*, 2nd ed. (Oxford: Oxford University Press, 2002).

to accept that, say, square pegs can fit into round holes. There can be structural constraints in a non-classical setting.

By accepting the law of non-contradiction one thereby accepts that the world is more constrained than it might otherwise have been. We acknowledge, however, that there is a way of understanding what it is to be a structural constraint that makes it natural to rule out the law of non-contradiction. If we think of "structure" solely in a geometrical sense, as seems right when interpreting the notion of a structural constraint in the peg/hole case or even the rings of Saturn case, then it can seem very odd to treat the law of non-contradiction as a structural constraint. For it is clearly not geometrical. This raises the question: in what sense of "structural constraint" does the law of non-contradiction qualify? Here is the beginning of an answer: think of structural constraints in terms of laws. The laws of physics place structural constraints on the kinds of events that can occur in our universe. In the case of extra-mathematical explanation, it is likely that there are laws that are at least partly mathematical in nature, and that play a similar role. So too in the extra-logical case: our universe is governed by certain logical laws, one of which is the law of non-contradiction. The way in which the law of non-contradiction serves as a structural constraint, then, is the way in which laws constrain more generally.

Casting the issue in terms of laws is fruitful. By proceeding in this fashion, it is possible to sketch a general theory of extra-logical explanation. According to the deductive-nomological (DN) account of explanation advanced by Hempel,[28] a proposition P explains a proposition Q just when Q is deducible from P plus (i) the laws of nature and (ii) any accidental facts needed to make the derivation stand up. As Baker shows, it is possible to formulate a DN account of extra-mathematical explanation.[29] So too it is possible—in principle at least—to formulate a DN account of extra-logical explanation. To do this for the time-travel case, simply derive the conclusion that Tim fails to kill his grandfather from the law of non-contradiction plus facts about Tim and Grandfather.

Of course, the DN account has fallen on hard times. It is more common these days to understand explanation in difference-making terms: P explains Q when P makes a difference to Q. Difference-making is, in turn, typically understood in counterfactual terms.[30] Accordingly, P makes a difference to Q when $\sim P \mathbin{\Box\!\!\rightarrow} \sim Q$. Can difference-making

[28] See Carl Hempel, *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science* (New York: Free Press, 1965).

[29] Baker, "Are There Genuine Mathematical Explanations," *op. cit.*, pp. 234–35.

[30] Though not universally, see Strevens, *Depth*, *op. cit.*

be used to understand extra-logical explanation? One might think not: to fit the time-travel case into a difference-making framework, the law of non-contradiction would need to make a difference to Tim's failure to kill Grandfather. This requires making sense of the following counterfactual: if the law of non-contradiction had been false, Tim would not have failed to kill Grandfather. It is tempting to think—following Lewis—that all such counterpossibles are trivially true.[31] While this yields the result that the law of non-contradiction makes a difference to Tim's failure to kill Grandfather (the counterpossible is true, after all), it would be a pyrrhic victory.

We should, however, resist the idea that all counterpossibles are trivially true. Recent work on counterpossibles has shown that standard semantic accounts of counterfactuals can be modified to allow for the non-trivial assignment of truth-values to counterpossible claims.[32] This includes cases such as the one above, where we are explicitly considering situations in which the law of non-contradiction fails. This general point about counterpossibles is connected to the point made above regarding non-classical logics: giving up the law of non-contradiction need not lead to explosion, and so there are worlds—albeit logically impossible ones—in which the law of non-contradiction fails and yet it is not the case that the world trivializes; it is not the case that everything (and its negation) is true in such worlds.

We often reason in non-trivial ways with counterpossibles of this kind. The grandfather paradox is one such case. Intuitively, a world with a brief glitch in the law of non-contradiction at exactly the moment of Tim's attempt on Grandfather's life such that, at the next moment, Grandfather dies is closer to the actual world than a world with a brief glitch in the law of non-contradiction at exactly the moment of Tim's attempt on Grandfather's life such that, at the next moment, Grandfather lives. So the counterpossible: had the law of non-contradiction been false, Tim would not have failed to kill Grandfather is true. The counterpossible intuitions that support this

[31] See David Lewis, *Counterfactuals* (Oxford: Blackwell, 1973), p. 24.

[32] See, for instance, Jc Beall and Bas van Fraassen, *Possibilites and Paradox: An Introduction to Modal and Many-Valued Logic* (Oxford: Oxford University Press, 2003); Berit Brogaard and Joe Salerno, "Remarks on Counterpossibles," *Synthese*, CXC, 4 (March 2013): 639–60; Edwin D. Mares, "Who's Afraid of Impossible Worlds?," *Notre Dame Journal of Formal Logic*, XXXVIII, 4 (September 1997): 516–26; Daniel Nolan, "Impossible Worlds: A Modest Approach," *Notre Dame Journal of Formal Logic*, XXXVIII, 4 (September 1997): 535–72; Graham Priest, "Sylvan's Box: A Short Story and Ten Morals," *Notre Dame Journal of Formal Logic*, XXXVIII, 4 (September 1997): 573–82; Greg Restall, "Ways Things Can't Be," *Notre Dame Journal of Formal Logic*, XXXVIII, 4 (September 1997): 583–96; David Ripley, "Structures and Circumstances: Two Ways to Fine-Grain Propositions," *Synthese*, LXXXIX, 1 (November 2012): 97–118.

result are just the standard intuitions about time travel: Tim has what it takes to kill grandfather, he is in the optimal situation to do so, and yet he fails: something gets in his way. So take away the principal barrier to Tim's success—namely the law of non-contradiction—and see what happens. On the face of it, Tim would (or at least might) succeed in killing his grandfather.

All of this, however, hangs on our claim that the law of non-contradiction is the best explanation of Tim's failure to kill Grandfather. One might deny this claim. According to Goddu, it is not logically impossible for Tim to kill Grandfather.[33] It is only an empirical impossibility, due to the nature of time in our universe, which is one-dimensional. If time were two-dimensional Tim would be able to kill Grandfather. For then moments of time would be divisible into distinct moments of hypertime and, in addition to being temporally extended, Tim and Grandfather would be hypertemporally extended. So Tim can kill Grandfather by preventing his Grandfather's hypertemporal parts from propogating any deeper into hypertime, without also killing him in time and inducing a paradox (see Figure 1).

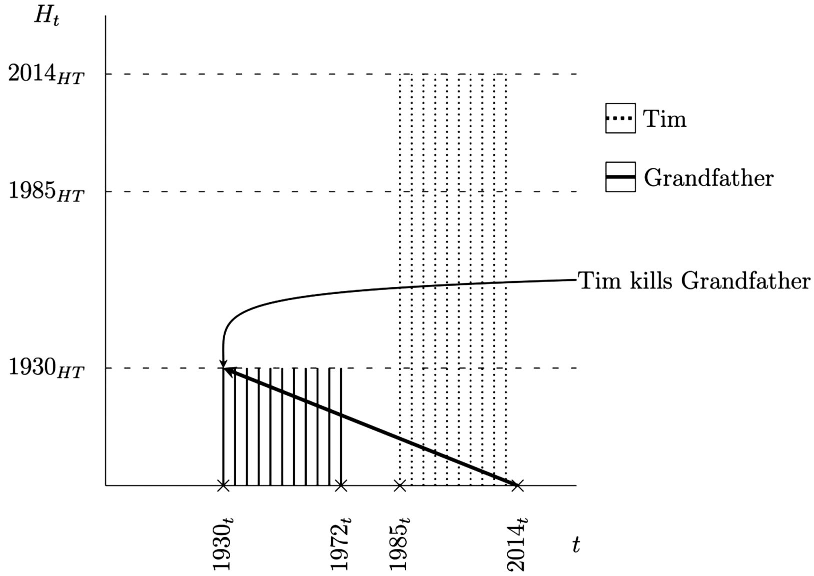

Figure 1. Time in Hypertime

[33] See G. C. Goddu, "Avoiding or Changing the Past?," *Pacific Philosophical Quarterly*, XCII, 1 (March 2011): 11–17; and G. C. Goddu, "Time Travel and Changing the Past: (Or How to Kill Yourself and Live to Tell the Tale)," *Ratio*, XVI, 1 (March 2003): 16–32.

Once it is conceded that Tim can kill Grandfather in hypertime a further explanation becomes available for why Tim fails to kill Grandfather actually. Tim fails because there is only one time dimension. This "dimensional" explanation sits at a higher level than the causal explanation, unifying Tim's various failures. It also has modal force: Tim cannot succeed so long as time is one-dimensional. But it does not appeal to the law of non-contradiction and, it could be argued, is all the better for it. So the dimensional explanation is the best explanation, not the extra-logical one.

Until now we have been restricting our interests to the Lewis case. Our aim has been to shed light on the kind of explanation that Lewis offered, by classifying it as a case of extra-logical explanation. We have thus been tacitly focusing on a fairly narrow explanandum: namely, Tim's failure to kill Grandfather *within his own timeline*. The dimensional explanation is no help here. For n time dimensions, Tim will fail to kill Grandfather within whichever temporal dimension he happens to call home. But perhaps it is the more unrestricted explanandum of why Tim fails *simpliciter* that is really at stake. For this case, it could be argued that appealing to a mid-level structural constraint such as the dimensionality of time will always be more illuminating than appealing to something as high-level and abstract as the law of non-contradiction.

While we can grant that, on its own, the law of non-contradiction is not particularly explanatory (at least, with regard to the unrestricted time-travel explanandum) it does not follow that the law of non-contradiction is explanatorily idle. For the law of non-contradiction to do the relevant explanatory work, we must combine it with (at least) the assumption that time is one-dimensional. The resulting explanation is not *purely* extra-logical—it is not a matter of logic alone explaining physical facts. But similar considerations apply to the extra-mathematical case: mathematics on its own may not explain the physical facts, but mathematics plus certain empirical assumptions often does; mathematics brings something extra, explanation-wise, to the table. So too, in the time-travel case. The law of non-contradiction adds something to the explanation for why Tim fails to kill Grandfather, something we cannot get from empirical assumptions about the nature of time alone. The logic is playing an indispensable explanatory role.

The law of non-contradiction, in particular, gives us two things. First, it furnishes the dimensional explanation with a greater modal scope. Think counterpossibly: for every world—be it possible or impossible—in which the law of non-contradiction holds and time is one-dimensional, Tim fails to kill Grandfather. But it is not the case that for every world in which time is one-dimensional, Tim fails to kill Grandfather. That is because, as noted above, there are impossible worlds in which a glitch in the logical laws permits Tim to kill Grandfather within his own timeline. The law of non-contradiction thus guarantees Tim's failure in a way that the empirical facts alone do not.

Second, the law of non-contradiction allows for a greater degree of explanatory unification. For within any hypertime world it is always possible to reformulate a paradox. Consider, for instance, hyper-Tim from 2014_{HT}. Just as Tim's existence is a downstream product of one of Grandfather's temporal parts, so too is hyper-Tim's existence a downstream product of Grandfather's hypertemporal parts. Grandfather lives until 1972_{HT}. Hyper-Tim—like Tim—wants to kill his grandfather. So he travels back through hypertime to 1930_{HT} and makes an attempt on Grandfather's life. If hyper-Tim succeeds then Grandfather's hypertemporal parts will never produce hyper-Tim, who will never travel back to make the attempt (see Figure 2).

Figure 2. Hypertime Paradox

What is the explanation for hyper-Tim's failure? We say: the law of non-contradiction. Goddu might say: the lack of a third temporal dimension. But for any world with n temporal dimensions, a paradox can always be formulated within a given dimension.[34]

[34] For a demonstration of this fact, see Sam Baron, "Back to the Unchanging Past," *Pacific Philosophical Quarterly*, forthcoming.

For each n-dimensional world, there will always be some dimensional explanation of why Tim or hyper-Tim or hyper-hyper-Tim fails to kill Grandfather, but the explanation will be different in each case, individuated by the particular dimension called upon. The law of non-contradiction, however, is available in every such case to explain why it is that n-dimensional Tim fails to kill n-dimensional Grandfather: it is the same explanation every time. The situation is similar to the rings of Saturn case. There is always some causal explanation for why it is that any given particle fails to orbit in the gaps in the rings of Saturn. However, a distinct causal explanation will be required for each case. The eigenanalysis ties all of these causal explanations together. So too in the time-travel case. The extra-logical explanation ties together all dimensional explanations providing a higher-level and ultimately more informative understanding of the situation.

In order to put to rest any lingering doubts about the logical explanation in the time-travel case, consider the following spatial analogue. Suppose you were in Sydney, Australia, at 9:00 am on September 20th, 2012. One thing that follows from this is that you were not somewhere else at this time. In particular, given that you were in Sydney, you were not in Helsinki, Finland. Moreover, any attempt to get to Helsinki by the time in question must have ended in failure. Now this failure can be made to sound mysterious, if we focus on the causal details: last-minute canceled flights, inexplicably oversleeping, repeated taxi breakdowns, slipping on banana peels, and so forth. But there is nothing mysterious here. Of course, if you did try to get to Helsinki by 9:00 am on September 20th, 2012, you must have failed, and there will be some particular story of your (perhaps repeated) failed attempt(s). In almost all contexts imaginable, the explanation for why you were not in Helsinki will be the causal one. But suppose there were an odd request for an explanation of why you were not in Helsinki *given that you were in Sydney*. In this case, you would need to invoke logic (and a bit of geography) in the explanation. But in the time-travel case, the request for an explanation of why Tim fails to kill his grandfather is like this odd request. The requested explanation is, in effect, of why Tim fails to kill his grandfather *given that his grandfather lives*. Once put like this, we see that Lewis is right to invoke logic at this point. Given that the time traveler was born the grandfather must have lived, and any attempted assassination by the time traveler, or by anyone else for that matter, must have failed.

I.4. Of a Piece. The eigenanalysis of the rings of Saturn is an instance of extra-mathematical explanation, and Lewis's solution to the grandfather paradox is an instance of extra-logical explanation. While both are independently interesting, what is of most interest to us is that the explanations are of a piece. In both cases there is a causal explanation available—a process explanation, if you like—as well as an explanation appealing to salient non-causal features that place structural constraints on physical systems—a program explanation, if you like. Moreover, in both cases there is a sense in which the causal explanation is inadequate because it fails to encode certain modal information, and in both cases the program explanation is deemed to be superior—for at least some purposes—precisely because it encodes this modal information, providing an explanation for why, in each case, the explanandum must be thus and so (more on this shortly). This parity between the two cases has the following upshot, to be discussed further in section II: taking the program explanation seriously in one case requires taking seriously the program explanation in the other.

II. DISCUSSION

The analogy between extra-mathematical and extra-logical explanation is an important one. While we find the connection between these two cases to be surprising in its own right, it has particular implications for three areas.[35]

II.1. Extra-Mathematical Explanation. As noted from the outset, some are skeptical of extra-mathematical explanations. Although this skepticism is most vivid in the debate over mathematical ontology, it runs beyond metaphysics. Skepticism about extra-mathematical explanation also arises in science.[36] Given the parity between cases of extra-mathematical explanation and cases of extra-logical explanation, however, this skepticism now comes at a greater cost.

As noted, extra-logical explanation and extra-mathematical explanation are of a kind, so we need to either admit them both or reject them both. In particular, the similarity between these two patterns of explanation suggests that extra-mathematical explanation is no more

[35] We should note that we are assuming that there is a difference between extra-logical and extra-mathematical explanations. For a logicist, of course, there is no distinction here: mathematics just is logic. For others, the distinction may be vague. We do not want to be drawn into a discussion of the boundary between logic and mathematics. Instead, we will rely on the accepted wisdom that there is some boundary or other between mathematics and logic. In any case, all we are committed to is that there are some explanations where the work is being done by what are commonly taken to be principles of logic (such as the law of non-contradiction in the time-travel case) and others where the explanation rests on mathematics and invoking mathematical objects (such as the number theory explanation of the cicadas).

[36] See, for example, the revealing exchange between Collin Rice, "Optimality Explanations," *op. cit.*, and Angela Potochnik, "Optimality Modeling and Explanatory Generality," *Philosophy of Science*, LXXIV, 5 (December 2007): 680–91. Even here, though, non-causal explanation is not without its supporters; see for example Elliot Sober, "Equilibrium Explanation," *Philosophical Studies*, XLIII, 2 (March 1983): 201–10.

or less objectionable than Lewis's account of the grandfather paradox. Hence, if the skeptic denies that mathematics is genuinely explanatory, then it seems she must also deny that logic is genuinely explanatory, at least in the time-travel case outlined here. This, in turn, would require rejecting the solution to the grandfather paradox that Lewis offered and the associated explanation for why it is that Tim cannot kill his grandfather. Conversely, if the skeptic of extra-mathematical explanation accepts Lewis's solution to the grandfather paradox, and thus accepts that logic can carry the explanatory load within an explanation, then it appears she must also accept the extra-mathematical explanations offered above and, in turn, concede that mathematics can be genuinely explanatory.

While, taken together, these two points do not show that skepticism about extra-mathematical explanation is unsustainable, it does show that the view is harder to maintain than previously thought. In order to make a case against extra-mathematical explanation it is not enough to consider mathematical cases only. One must be prepared to deal with, at the very least, logical cases as well. We anticipate two responses to this line of thought.

First, one might take umbrage with the analogy between mathematical cases, such as the rings of Saturn case, and logical cases, like the time-travel case. In particular, one might argue that while mathematics is not genuinely explanatory, logic is, thereby preserving skepticism about extra-mathematical explanation while retaining a commitment to Lewis's solution to the grandfather paradox. However, it is not obvious what the point of disanalogy might be. As we have argued, both kinds of cases are non-causal, both involve structural constraints on physical systems, and both have a robust modal character. The only difference seems to be that one is logical and one is mathematical, but it is hard to see why that should be enough to ground the claim that logic, and not mathematics, can be explanatory. Indeed, if anything, one would expect the opposite result. While the explanatory contributions of mathematics are increasingly well documented, few cases in which logic is genuinely explanatory have been identified to date. Hence, the more plausible position appears to be that mathematics, and not logic, is explanatory. But this will not help the skeptic's position. So, again, more is required of the skeptic: she must try to pull apart the cases identified here in a manner that suits her purposes, and this looks hard to do.

Second, the skeptic might simply bite the bullet and reject Lewis's solution to the grandfather paradox, thereby leaving herself room to deny that mathematics is genuinely explanatory in the rings of Saturn case. For such skeptics, the only explanation available of Tim's continual failed assassination attempts is the sum of the various causal details: the gun jammings, the changes of heart, and so on. However, as has been noted, these apparently unrelated causal histories eventually mount up to look so implausible as to suggest that there is a conspiracy in place. It is for this reason that all parties to the debate over time travel agree that, on its own, the causal explanation Lewis offered is insufficient. Hence, denying that the extra-logical explanation of why Tim cannot kill his grandfather is genuinely explanatory comes at a significant cost. One must go on to either address the criticisms of the causal explanation that Lewis offered, and that continue to fuel the debate over time travel, or provide a new explanation for why Tim cannot kill his grandfather, one that is not of a piece with mathematical explanation but which nevertheless unifies Tim's various failures. To our knowledge, no such explanation has, as yet, been advanced.

Thus, all is not smooth sailing for the skeptic of extra-mathematical explanation. She must either defeat the analogy offered here between extra-logical and extra-mathematical explanation or she must provide a new solution to the grandfather paradox. While we do not rule out either option, each is a substantial project. Moreover, each is a project that skeptics of extra-mathematical explanation have seen no need to embark on up until now. Given what we have said here, this is a substantial oversight. Extra-mathematical explanation does not stand alone. Such explanation is a member of a more general class of noncausal explanations, and so it is with the class and not the instance that the skeptic of extra-mathematical explanation must take issue.

II.2. Time Travel. Until now, Lewis's solution to the grandfather paradox has seemed to be one of a kind. But, in fact, Lewis's logical solution to the grandfather paradox is in good company. The explanation he offered is really no different in kind from extra-mathematical explanations and, indeed, from a great many scientific explanations that are geared toward answering high-level "why" questions. So if one rejects Lewis's solution, as some are inclined to do, then one incurs an extra cost. This is the reverse of the point made in section II.1: just as skepticism about extra-mathematical explanation is now more costly, because it requires taking a stand on time travel; so too is skepticism about Lewis's solution to the grandfather paradox more costly, because it requires taking a stand on extra-mathematical explanation.

When resisting this line of thought, then, there are similar options open to the skeptic of extra-logical explanation as there were open to the skeptic of extra-mathematical explanation. First, one might take issue with the analogy between the two cases. As before, however, it is difficult to see what the point of disanalogy might be. Second, one might deny that logic is explanatory but accept that (or remain neutral on whether) mathematics is explanatory, thereby avoiding the need to take a stand on the debate over extra-mathematical explanation. As noted, this option might seem tempting: there are few cases of logical explanation documented thus far, and so perhaps skepticism about extra-logical explanation is easier to sustain. One might press the point further by arguing that if we accept that logic is explanatory in the time-travel case, then explanations of this kind are far too easy to come by. Why is it not both raining and not raining? Because of the law of non-contradiction. Arguably, such trivial cases are analogous to the logical explanation appealed to in the grandfather paradox and yet, surely, we would not want to count them as genuine explanations, or so this line of argument would run.

Actually, we are not so quick to reject such cases as non-explanations. They are perhaps uninteresting explanations or answers to "why" questions that are never asked. Nevertheless, we recognize the countervailing intuition that they are not genuinely explanatory. This intuition, however, can be explained away by thinking about the contextual element of explanation. It is clear that some explanations seem more salient than others, depending on one's background context. This notion of salience is somewhat vague, but should be familiar enough: it is, for instance, the reason why sometimes all we care about is the causal explanation, while other times we care more about some non-causal explanation in the neighborhood. What is interesting about extra-logical explanation is that there are very few contexts in which a logical explanation is the kind of explanation that one is interested in, and so very few contexts in which these explanations appear salient. This, we contend, is why there are so few documented cases of extra-logical explanation and thus why skepticism about this kind of explanation can seem so tempting. But if that is right, then the temptation should be resisted: we should not reject a potential class of explanations simply because there are very few instances of that class that we actually care about.

Perhaps we are being too generous to ourselves. It might be thought that the space of extra-logical explanations is so sparsely populated as to be empty. This may be right for *purely* extra-logical explanations: cases in which the explanatory load is carried by logic alone. Things are different, however, when we move away from pure extra-logical explanations, to cases in which logic is playing an indispensable explanatory role within a broader explanation containing non-logical elements. Explanations of that kind are, we believe, much easier to locate. Indeed, in many of the cases of extra-mathematical explanation offered to date, we find consistency constraints playing a role. To be sure, consistency is not the whole story: mathematics is in there too. But in some such cases, if one were to strip out the law of non-contradiction, the explanation might well fall over.[37]

So it may be that extra-logical explanations are widespread indeed: wherever consistency is boosting the explanatory credentials of a theory, logic is likely to be doing some explanatory work. Still, due to the aforementioned contextual element, we rarely have reason to focus on the consistency constraints in these explanations. The time-travel case is thus very special: the possibility of time travel opens up a context in which logical explanations involving the law of non-contradiction become very salient indeed.[38]

Any further sense one might have that Lewis failed to explain why Tim fails to kill his grandfather is, we suspect, a result of attending too closely to the wrong "why" question. If one attends to the causal "why" question—that is, the low-level question of what it is that, as a matter of fact, explains why Tim fails—then it can seem as though Lewis missed the point and thus failed to solve the paradox. If we attend primarily to the higher-level question, however, then Lewis's solution is perfectly adequate. Now, in the face of Lewis's solution, you can certainly demand a low-level explanation for the high-level question of why Tim fails to kill his grandfather, but causal explanations do not supply the right kind of information to answer the high-level question. A similar idea seems plausible in the extra-mathematical case. In so far as some remain skeptical of explanation of this kind, such skepticism may be the result of too closely attending to low-level "why" questions. Again, in the face of an extra-mathematical explanation, one can certainly demand a low-level answer to a high-level question, but that would be to make the same sort of mistake.

II.3. Explanation. We come, at last, to theories of explanation more generally. As noted briefly in the introduction, there are those who seem to be committed to the view that causal explanation is all there is. Although we suspect that such a view of explanation is quite widely

[37] There is also an interesting question about how much logic is required for the various candidates for mathematical explanations found in the literature. Perhaps some of the so-called mathematical explanations might be better thought of as logical explanations. To pursue such a line, however, would require taking a stand on the thorny issue of the boundary between logic and mathematics. As we have already noted, we would like to avoid that issue here. For present purposes, all we require is that eigenanalysis is a part of mathematics and the law of non-contradiction is a part of logic.

[38] One might worry here: if the only way to explain why Tim fails to kill Grandfather is to appeal either to logic (which, for some, seems odd) or to causation (which is unsatisfactory) then perhaps the moral to be drawn is that time travel is unintelligible, despite making grammatical sense. This would be a strong claim, however, since time travel is recognized to be a genuine physical possibility. We do not wish to enter into such issues here and simply assume that there is evidence enough to take time travel to be intelligible.

held, it is, with some notable exceptions, difficult to find philosophers endorsing this general position in print. What we find, instead, is that philosophers interested in explanation tend to accept that there might be non-causal explanations, but that such cases (if any there be) are beyond the scope of the particular investigations in which they are engaged.[39] Their investigations are, rather, geared only toward offering an account of causal explanation and not, necessarily, an account of scientific explanation in general. Implicit in this approach, however, and its associated hedging, is the idea that causal explanation is, in some sense, where the action is. Non-causal explanation is, by contrast, deemed to be of lesser importance, and as such a philosophical exploration of explanation can legitimately restrict its purview to the causal case, without fear of restricting too far or of leaving out too much.

But a theory of causal explanation has trouble with the cases we have considered here. While a causal theory of explanation can capture the causal processes that arise in the time-travel case, on its own the low-level causal explanation for why Tim fails to kill his grandfather is generally thought to be inadequate, and it is not obvious that any explanation along these lines will do better. Similarly, causal explanations come up short when considering various stability phenomena such as the gaps in the rings of Saturn. This is, perhaps, unsurprising; in both cases the important question appears to be the high-level "why" question. What we really want to know, then, is whether a suitably sophisticated theory of causal explanation can be extended to cover the high-level explanations in these cases, treating them as special kinds of high-level causal explanations. If so, then the restriction to causal explanations may be acceptable, *modulo* the discovery of further recalcitrant cases. If not, then there is pressure on those engaged in the philosophy of explanation to develop new frameworks in which to capture these high-level explanations.

It is clear that some causal theories of explanation, such as the process theory defended by Salmon, cannot be extended in this way. That is because, at this higher level, there are no causal processes that might be used to underpin the relevant explanations. A causal process, for Salmon at least, is the intersection of two world-lines in space-time, across which some conserved quantity (energy, momentum, and so on) is transferred. In this sense, there is no causal process connecting the laws of logic with Tim's various failures to kill his grandfather. It is not even clear that this makes sense. Similar considerations apply to the rings of Saturn case.

[39] See, for instance, Wesley Salmon, *Scientific Explanation and the Causal Structure of the World* (Princeton: Princeton University Press, 1984); Strevens, *Depth*, *op. cit.*; and Woodward, *Making Things Happen*, *op. cit.*, who hedge in this way.

The more nuanced theories of causal explanation developed by Woodward and Strevens, by contrast, may do better. Broadly speaking, these accounts seek to understand causal explanation in terms of difference-making, rather than causal processes. Difference-making, for Woodward at least, is to be understood via classes of counterfactual dependencies, modeled structurally.[40] The difference-making framework is rather more flexible than Salmon's process account, since it deploys a more liberal conception of causation, and so it may be possible to extend it to the kinds of high-level explanation in which we are interested. Indeed, we suspect that some possess an optimistic tendency toward thinking that this can be done with relative ease and thus that all explanations, even high-level ones, can be ultimately subsumed within a causal difference-making framework.[41]

By attending to the analogy between the extra-logical and extra-mathematical cases, we can begin to see the difficulty with attempting to extend difference-making accounts in this way and therefore why a causal theory of explanation cannot be the whole (or even most of the) story about explanation. As touched on briefly above, one important similarity between the time-travel case and the rings of Saturn case is that the high-level "why" questions in each case are requesting modal information; information that is not supplied by any low-level instances of causation. In the time-travel case, the low-level causal explanation does not tell us why Tim cannot kill his grandfather, no matter how hard he tries, and yet that is what we want to know. Similarly, in the rings of Saturn case, an important modal component of the explanation is provided by the mathematics but the low-level causal explanation misses this. That the high-level explanation requires some modal information does not prevent it from being situated within a causal framework. Difference-making accounts of the Woodward variety do deal in the modality of explanation, up to a point. Because these accounts rely on counterfactuals, explanations modeled in this way typically possess some modal force. But these counterfactual accounts, at least as they are standardly developed, do not imbue explanation with a modal force strong enough to cover the cases considered here.[42]

[40] For Strevens, difference-making is understood in terms of causal entailment. Roughly: x makes a difference to y iff for some argument A with premises $P_1 \ldots P_n$ and conclusion y, the P_n plus x causally entail y whereas the P_n minus x do not causally entail y; where ψ causally entails φ when the derivation $\psi \vdash \varphi$ represents an actual causal process that has φ as its consequence, and where $\psi \vdash \varphi$ represents an actual causal process when each premise in ψ uses modus ponens to deduce the next premise from the causal laws plus initial conditions.

[41] See, for instance, Skow, "Are There Non-Causal Explanations," *op. cit.*

[42] The difference-making account defended by Strevens is also modal in character. Because it operates in terms of entailment, it has the same resources available to it as

We can see this by considering, again, the difference-making account of logical explanation sketched in section I.3. If the law of non-contradiction genuinely makes a difference in the time-travel case, then we must be able to make sense of counterfactuals involving the laws of logic, such as the following:

(CF_1) If the law of non-contradiction had been false, then Tim the time traveler might have been able to kill his grandfather.

Similarly, in the rings of Saturn case, we must be able to make sense of counterfactuals involving the eigenanalysis of the system, such as:

(CF_2) If the mathematical eigenanalysis of the Saturn system had been different, the gaps in the rings of Saturn might have been different.[43]

Because we are interested in the extent to which the law of non-contradiction or the eigenanalysis of a planetary system makes a difference to how things are in the physical universe, we evaluate these counterfactuals by holding fixed everything within the explanation except for the law of non-contradiction, or the eigenanalysis, and then attempt to determine what effect there would have been on the explananda had these things been different. As we have already seen, evaluating the likes of CF_1 and CF_2 will require something other than standard counterfactual machinery—perhaps some paraconsistent analogue. For CF_1, we must consider what would have happened in some scenario in which the law of non-contradiction is false, and there are true contradictions. For CF_2, assuming that mathematical truths are necessary truths, evaluating what would have been the case had the eigenanalysis worked out differently takes us into an impossible scenario (there is only one way that the analysis can work out, and necessarily it works out this way), and thus again into the realm of assessing counterfactuals with impossible antecedents.

For some, the mere fact that generalizing the difference-making account in this way moves us into non-classical terrain would count as a strong reason against any such generalization. But we are willing to be liberal about what can make a difference, and so we are prepared to allow difference-making of the relevant kind across both possible and

---

the DN account of explanation. As discussed above, it is possible to develop a DN-style account that can handle logical and mathematical explanation. So it may be possible to generalize the Strevens picture as well. The same issues that arise for the generalization of Woodward's account will arise here too: Strevens's account uses causal entailment, and such a thing does not sit well with logical and mathematical explanations.

[43] For an application of counterfactuals to cases of extra-mathematical explanation see Sam Baron, "The Explanatory Dispensability of Idealizations" *Synthese*, forthcoming.

impossible worlds. What we resist, however, is the idea that once the difference-making account is extended in this way, the relevant high-level explanations count as *causal explanations*. Once difference-making has been extended to include counterpossible cases, it is too far removed from what we usually mean by 'causation' to count as an analysis of causal explanation. Causation is typically thought to be underpinned by the actual laws of nature. On a difference-making account, the evaluation of standard counterfactuals usually makes use of a closeness metric between possible worlds, one that relies heavily on nomic similarity. The kinds of cases we are considering, however, are not constrained by the laws of nature in the right way to be able to deploy the same closeness metric. Because we are dealing with counterpossible cases, we are automatically dealing with counterfactuals that to be evaluated must be situated within a much more general framework than that of nomic possibility.

The upshot, then, is this: while it may be an option to extend a difference-making conception of causal explanation to cover the high-level mathematical and logical explanations we have considered here, doing so severs the conceptual connection with causation. To press this point, it is worth considering a somewhat different case: the case of intra-mathematical explanation (the explanation of one mathematical fact by another). For example, a mathematical theorem might be explained by a particular proof. For instance, consider the obvious proof by induction that all natural numbers of the form $2n+1$ (where n is a natural number) are odd. Such a proof proceeds in two steps. First, we show that the base case holds; then we prove the inductive step, that if it holds for some number k it also holds for $k+1$.

Arguably, the inductive proof explains why it is that for all natural numbers n, $2n+1$ is odd; actually this is controversial,[44] but let us grant it for now.[45] What we are interested in, then, is the relationship between explanations of this kind and the generalized version of the difference-making account just considered. Let us suppose that the difference-making account is generalized to handle counterpossible cases, such as cases in which mathematical facts are assumed to be false, in order to determine the difference this makes to an explanation. If one is already prepared to extend the difference-making account that far, then it can be pushed a bit further to cover the intra-mathematical case as well. That is, we can model the nature of intra-mathematical explanation via difference-making by considering

[44] See, for instance, Marc Lange, "Why Proofs by Mathematical Induction Are Generally Not Explanatory," *Analysis*, LXIX, 2 (April 2009): 203–11; and Alan Baker, "Mathematical Induction and Explanation," *Analysis*, LXX, 4 (October 2010): 681–89.

[45] If you do not like this example, choose your own favorite case of a mathematical theorem explained by a particular proof.

what differences would result from manipulating the mathematics within a proof. If, for example, it was not the case that $2n+1 = 3$ where $n = 1$, but, rather, $2n+1 = 4$ where $n = 1$, then the base case in the above proof would fail and thus so would the proof. We would then no longer have an explanation of the inductive conclusion. Of course, it is impossible for $n = 1$ and for $2n+1 \neq 3$ (assuming that what we mean by each of these symbols is held fixed), and so to model the difference that the base case makes to the rest of the proof we need to invoke counterpossibles. The point, however, is that there is nothing special about this kind of case, compared to the other instances of counterpossibles needed to model, say, the law of non-contradiction. If one already allows for non-classical difference-making of this kind, then it would seem one opens the door to difference-making accounts of intra-mathematical explanation. Indeed, it is hard to resist the push to extend the account in this way. After all, doing so yields an incredibly unified and robust account of explanation. Whatever one thinks of such an account, however, it is clear that difference-making generalized into intra-mathematical explanation would not be a causal account. The base case is not *causing* the result; that is absurd.

The general moral, then, is this: difference-making in a non-classical scenario does not line up, naturally, with causation. The modal character of high-level explanations of the kind considered here is therefore much stronger than the modal character of causation, and so cannot be easily subsumed under a theory of causal explanation. This point dovetails with recent work by Lange,[46] who also argues that the modal character of a mathematical explanation is stronger than any causal explanation. Because extra-mathematical explanations typically describe a framework within which any possible causal relation can be situated, argues Lange, including some that are beyond the scope of nomic possibility, the modal character of such explanations is stronger than the strongest conception of causation (cashed out in terms of laws). Although Lange does not consider extra-logical explanations, his view generalizes to those explanations as well. As with extra-mathematical explanations, extra-logical explanations operate against a framework within which all possible causal relations can be situated. This framework can then be used to provide explanations for physical phenomena that outrun any explanations involving causation. This imbues extra-logical explanations with a similar kind of necessity as that of extra-mathematical explanations and, again, a kind of necessity that cannot be found in even the strongest causal explanation. So

[46] See Marc Lange, "What Makes a Scientific Explanation Distinctively Mathematical?," *The British Journal for the Philosophy of Science*, LXIV, 3 (September 2013): 485–511.

the diagnosis offered above remains apt: theories of explanation require something more than causation in order to accommodate the modal character of high-level explanations. The analogy between the extra-logical and extra-mathematical cases helps us to see that.

III. CONCLUSION

This paper might be thought of as a plea for philosophers to look beyond the confines of causal explanation, even when that notion is cashed out in liberal, difference-making terms. Embracing non-causal explanation as a useful supplement to causal explanation gives us greater explanatory resources and would seem to square better with at least some scientific practice.[47] To be sure, causal explanation (or process explanation) is an important class of explanation, and in many scientific contexts it may be all that is required. But there are other contexts where non-causal explanation is the more informative. Apart from anything else, looking beyond causal explanations will allow one to see past the conspiracy of failed assassination attempts in the grandfather paradox and appreciate the power of the Lewis solution. It will also allow one to see the power of the Arnol'd account of Saturn's rings and other instances of mathematical explanation. Our plea is not new. Others have drawn attention to non-causal explanation in other contexts.[48] Our modest contribution here is to note, on the one hand, that the debate over extra-mathematical explanation is bigger than the philosophy of mathematics, and, on the other hand, that the debate over Lewis's solution to the grandfather paradox is bigger than the philosophy of time. Both cases are part of a larger class of non-causal explanations and, as such, are involved in a much broader debate about the nature of explanation. This makes it much harder to take a stand on any one of these issues in isolation and, moreover, helps us better see the direction that the philosophy of explanation must take.

SAM BARON

University of Western Australia

MARK COLYVAN

University of Sydney

[47] After all, the eigenanalysis and similar stability selection accounts are advanced as explanations in various branches of science. There is also the issue of accounting for intra-mathematical explanation, which is an accepted part of mathematical practice; see chapter five of Mark Colyvan, *An Introduction to the Philosophy of Mathematics* (Cambridge, UK: Cambridge University Press, 2012).

[48] Such as Smart, "Explanation – Opening Address," *op. cit.*; Steiner, "Mathematics, Explanation, and Scientific Knowledge," *op. cit.*; Steiner, "Mathematical Explanation," *op. cit.*; and Graham Nerlich, *The Shape of Space*, 2nd ed. (Cambridge, UK: Cambridge University Press, 1994).