

BOOK REVIEWS

Causation and Free Will. CAROLINA SARTORIO. Oxford: Oxford University Press, 2016. viii +200 pages. Cloth \$65.00.

The problem of free will has lost a bit of its metaphysical edge in recent years and is due for a sharpening. Seen in this light, Carolina Sartorio's recent book is a 1,000-grit whetstone. It is a fresh and innovative take on a familiar set of problems, one that clearly demonstrates the theoretical advantages of taking metaphysics seriously. Not only does Sartorio achieve her modest goal of showing that "the concept of causation. . . plays a central role in an attractive solution to the problem" of determinism and free will (3), but she also, in my view, makes a convincing case for the conclusion that if you are working on the metaphysics of free will but are not thinking seriously about the nature of the causal relation, then you are probably doing it wrong.

The book is a short and relatively easy read, with three substantive chapters bookended by material that readers familiar with the state of the debate can safely skim. In chapters 2 and 3, Sartorio argues for several controversial claims about causation that, if adopted, can provide the foundations for a view of free will according to which it is wholly grounded in facts about the actual causal history of the free action—and not, for example, in facts indicating whether an agent was able to do otherwise. The idea that freedom is best understood as an *actual-sequence* notion (rather than *alternative-sequence* notion) is familiar from the work of Frankfurt and Fischer, among others.¹ But Sartorio maintains that her underlying views about the causal relation provide the best foundation for such an account. In particular, Sartorio argues that the best way to conceive of the *actual-sequence* is as the actual *causal* sequence, and the best way to conceive of the causal relation is as an intransitive, extrinsic, difference-making relation that can take both *absences* and positive events as relata. Chapters 2 and 3 develop and defend the idea that if this is what the causal relation looks like, then compatibilists need not look beyond the actual causes of an action to see why it counts as free.

In chapter 4, Sartorio pushes the idea even further: not only do the actual causes of an action do as well as the infamous ability to do

¹ Harry G. Frankfurt, "Alternate Possibilities and Moral Responsibility," this JOURNAL, LXVI, 23 (Dec. 4, 1969): 829–39; John Martin Fischer, *The Metaphysics of Free Will* (Oxford: Blackwell, 1994).

otherwise when it comes to explaining our intuitions in a number of puzzling cases, but they can also help to explicate a widely accepted necessary condition on freedom and responsibility, namely the idea that an agent must be sensitive to an appropriately wide range of reasons. At first this requirement seems to clash with an actual-sequence account of freedom, since the most well-developed account of reasons-sensitivity is spelled out in terms of how an agent *would* react if they *were* to have or recognize reasons of a certain sort.² In contrast, Sartorio's ingenious idea is that even if counterfactuals like that are reliable indicators of reasons-sensitivity, what they indicate is that the actual causal history of the action includes an array of *absences* of reasons, and it is in virtue of the causal role played by these (actual) absences that the agent counts as reasons-sensitive.

In fact, the book as a whole might be described as a sort of refurbishment or renovation of the range of compatibilist views that have been inspired largely by the work of Frankfurt and Fischer. Many contemporary theorists accept actual-sequence accounts that include a requirement of reasons-sensitivity, but what Sartorio shows is that nobody really has had a clear idea of what 'actual sequence' even means, or of how to understand reasons-sensitivity in a way that is consistent with that. As Sartorio herself puts it, her book offers "a revitalized version of an actual-sequence view that is buttressed by a robust and well-motivated metaphysics" (44).

Let us get into some of the details. Thought experiments play a large role in the argumentative structure of Sartorio's book, and two in particular seem most central. First are the Frankfurt examples (which provide the original motivation for pursuing an actual-sequence account), but another sort of case to which Sartorio repeatedly returns is *All Roads Lead to Rome* (57) and its variants, in which Ryder is on a runaway horse approaching a crossroads. Ryder cannot stop the horse but he can steer the horse onto one of the forks rather than the other, and since Ryder wants to harm some Romans, he steers the horse onto the fork that Ryder thinks leads to Rome. In the "all roads" variant, although Ryder does not realize it, both forks lead to Rome, so Romans would have been harmed in any case. Over the course of the book, Sartorio compares this case to a similar case (*Not All Roads Lead to Rome*), which is exactly the same except that only Ryder's chosen fork leads to Rome. Cases like these are meant to illustrate the *extrinsic*, *intransitive*, and *difference-making* nature of the causal relation.

²This is a distressingly simplified description of the view found in John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge, UK: Cambridge University Press, 1998).

Start with difference-making. According to Sartorio, causes are difference-makers in the sense that “their effects wouldn’t have been caused by the absence of their causes” (94). Sartorio thinks that intuitively Ryder is not responsible for the harm to the Romans in the case where all roads lead to Rome, whereas he is responsible for the harm in the case where not all roads lead to Rome. This might naturally be explained by saying that when all roads lead to Rome, Ryder could not have avoided harming the Romans. But if causes are difference-makers, we can offer a purely actual-sequence explanation, namely: when all roads lead to Rome, Ryder’s act of steering does not cause any Romans to be harmed, since the same harm would have been caused by his not steering (and hence his act of steering does not make a difference in the sense relevant to the causal relation). This is not an *argument* for the difference-making constraint, of course, and as far as I can tell there is not much of an argument in the book for this constraint.³ (In fact, one of the primary ways Sartorio motivates the difference-making constraint here in the book is by asking the reader to reflect on thought experiments involving moral responsibility. But it is not clear that our judgments about moral responsibility are independent of our judgments about what causes what, so this argumentative strategy makes me a bit uneasy.⁴)

If we grant difference-making then extrinsicness follows, and it is a fairly short jump to intransitivity, as well. To say that the causal relation is extrinsic is to say that “a causal relation between C and E may obtain, in part, owing to factors that are extrinsic to the causal process linking C and E” (71). Again, take the case of Ryder. Whether the road not taken leads to Rome partly determines whether Ryder’s act of steering causes harm to the Romans; hence, the causal relation is extrinsic. And we can get intransitivity with minimal additional assumptions as follows: if all roads lead to Rome, then although Ryder’s act of steering causes the horse to go down a particular path, and (arguably) the horse’s going down that particular path causes the Romans to be harmed, Ryder’s act of steering does not cause the Romans to be harmed.

So far so good, but how does all of this help with developing an actual-sequence account of free will? Well, if facts about freedom are wholly grounded in actual causal facts, then any time someone is inclined to use “they couldn’t do otherwise” as an explanation of the lack of freedom, an actual-sequence theorist like Sartorio needs to

³But see Carolina Sartorio, “Causes as Difference-Makers,” *Philosophical Studies*, CXXIII, 1/2 (March 2005): 71–96.

⁴See, for example, Joshua Knobe, “Intentional Action in Folk Psychology: An Experimental Investigation,” *Philosophical Psychology*, XVI (2003): 309–24.

offer an alternative explanation that does not appeal to the ability to do otherwise. As I mentioned above, Ryder does not freely harm the Romans in *All Roads Lead to Rome*, and Sartorio can explain that by appealing only to the actual sequence: his act of steering does not even cause the Romans to be harmed. And with respect to Frankfurt cases: Sartorio can explain why the agent is free despite lacking the ability to do otherwise. The agent's action is caused by their own reasons, despite the fact that the action itself is inevitable. It is true that the absence of those reasons would have caused the intervener to intervene, which would have caused the agent to act in the same way, but the actual reasons can still be difference-makers, since the casual relation is not transitive.

So, the overall picture from chapters 2 and 3 is as follows: a close examination of the causal relation reveals the resources for a straightforward actual-sequence view of free will, according to which freedom is grounded solely in the actual causal history of the action. It is just that difference-making and intransitivity make reconstructing that causal history a more complicated affair than it at first might have seemed.

The account of reasons-sensitivity in chapter 4 is perhaps the most innovative part of the book, and I think it is worthy of serious consideration. The notion of "sensitivity" seems clearly to be a modal notion, and accordingly those accounts of reasons-sensitivity that are best worked out (for example, that of Fischer and Ravizza) invoke counterfactuals about how agents would behave were they to recognize certain reasons, or claims about the range of possible worlds in which an agent successfully recognizes the reasons they have, and so on. Even if such accounts are not meant to be reductive, Sartorio is right to demand an account that sits more easily within an actual-sequence framework. What distinguishes an addict from a non-addict is that the latter but not the former is appropriately reasons-sensitive when they take the drug, and an actual-sequence account should be able to cash that out by saying how the relevant causal histories differ, even if they both end with taking the drug. Sartorio does this by appealing to *absences*.

To oversimplify: the addict's act of taking the drug is not caused by the absence of sufficient reasons to do otherwise (as indicated by the truth of the counterfactual that if there were sufficient reasons to do otherwise, the addict would still take the drug); the non-addict's act of taking the drug *is* partly caused by the absence of sufficient reasons to do otherwise. Hence, the non-addict takes the drug freely and the addict takes the drug unfreely, and this is accounted for wholly by facts about the actual causes of the act of taking the drug. Responsiveness is not about what reasons you *would* respond to; rather, it is about what

absences of reasons you *do* respond to (that is, which absences are part of your action's actual causal history).

I am sympathetic to a lot of what Sartorio says in the book, but I do have two worries that are perhaps worth mentioning. First, I worry that Sartorio has pushed down one bump in the rug only to have another appear elsewhere. One of her major criticisms of extant actual-sequence accounts of reasons-sensitivity is that they go beyond the actual sequences in relying on counterfactuals. One of the virtues of her own view is supposed to be that "it avoid counterfactuals altogether" (134). But it is not clear that this is right. It is true that her account of reasons-sensitivity is not stated explicitly in terms of counterfactuals, and is instead a claim about the role that certain absences play in the actual causal history of free action. But earlier in the book she makes clear that causation requires difference-making, where difference-making is spelled out as follows: "Causes make a difference to their effects in that the effects wouldn't have been caused by the absence of their causes" (94). So, if reasons-sensitivity is a matter of certain absences playing a causal role, and if playing a causal role requires difference making, then reasons-sensitivity is ultimately to be understood in terms of whether the absence of certain absences would have had the same effect as the absences themselves.

That is a lot of tokens of the word 'absence', so let me make this a bit more concrete. According to Sartorio, when a non-addict's act of taking a drug is a display of reasons-sensitivity, this is because the act of taking the drug is caused, in part, by the absence of a sufficient reason to refrain. But now, if the causal relation is a difference-making relation, then that means that the absence of a sufficient reason to refrain could only cause the act of taking the drug if the absence of *that* absence would not also have caused the act of taking the drug. Or, in other words, the absence of a sufficient reason to refrain could only cause the act of taking the drug if *having* a sufficient reason to refrain would not have caused the agent to take the drug. But now that just looks like the more standard counterfactual account of reasons-sensitivity: the agent manifests reasons-sensitivity only if having a sufficient reason to refrain would not also have led the agent to take the drug. So it is not entirely clear how much genuine progress has been made by the move to absence causation.

My second worry is, I think, related. Suppose we accept the claim that whenever an action displays an agent's reasons-sensitivity, the actual causal history of the action will include some crucial absences of reasons. Every time I drive my usual route to the office, for example, I do many things that display my sensitivity to reasons. On Sartorio's account (127), I turn right out of my driveway partly because of the absence of cars

blocking that route. If that were not so—that is, if the absence of that absence, or the *presence*, of cars blocking that route equally would have caused me to turn in that direction—then I would not count as reasons-sensitive in the relevant respect. (This is because the absence of cars would not satisfy the difference-making constraint, and hence would not be part of the causal history of my choice to turn right, and hence would not be part of the story of my reasons-sensitivity.) But is the role that this absence plays in the causal history of my action part of *what it is* for me to be reasons-sensitive, or is it, rather, a mere *indication* of my reasons-sensitivity? And if it is only the latter, then we are still left wondering about the right account of reasons-sensitivity, and whether it can avoid appealing to counterfactuals.

To state the worry another way: there seems to be a difference between *being reasons-sensitive*, on the one hand, and *acting in a way that manifests one's reasons-sensitivity*, on the other. When a non-addict takes the drug, their action is partly caused by the absence of sufficient reason to refrain, and it is that particular causal history that makes the non-addict's reasons-sensitivity *manifest* on this occasion. But it seems odd at best to say that the obtaining of a causal relation between the absence of sufficient reason to refrain and the act of taking the drug is *what it is* for the agent to be reasons-sensitive. Rather, the particular causal history is what it is *because* the agent *is already* reasons-sensitive and displays that sensitivity on this occasion. Or at least so it seems to me. But perhaps I am simply revealing my anti-reductionist impulses.

Let me close by reiterating that this is an excellent book, sure to be of interest to anyone with a stake in the causation or free-will literatures. Furthermore, it nicely represents a welcome trend in philosophy, in which authors encourage us to confront complexity and structure in the actual world head on, rather than look for otherworldly proxies for or indications of it.⁵

NEAL A. TOGNAZZINI

Western Washington University

⁵For helpful conversations and written comments, thanks to Christopher Franklin and Justin Coates.