



# Causation and Responsibility

Carolina Sartorio\*

University of Wisconsin at Madison

---

## Abstract

In this article I examine the relation between causation and moral responsibility. I distinguish four possible views about that relation. One is the standard view: the view that an agent's moral responsibility for an outcome requires, and is grounded in, the agent's causal responsibility for it. I discuss several challenges to the standard view, which motivate the three remaining views. The final view – the view I argue for – is that causation is the vehicle of transmission of moral responsibility. According to this view, although moral responsibility does not require causation, causation still grounds moral responsibility.

---

## 1. Introduction

Agents can be responsible for things in different ways. In particular, they can be causally responsible for things, and they can be morally responsible for things (other notions of responsibility include, notably, legal responsibility, which will not be my focus here). The concepts of causation and moral responsibility seem to be importantly related to each other. Philosophers have discussed the connection between them from different perspectives: the perspective of action theory, ethics, and metaphysics. In this article I examine different proposals on the nature of that link.

Let us first be clear about our focus. It is natural to draw a distinction between the acts that an agent performs (his actions or omissions) and how things turn out in the world as a result of those acts (the events and states of affairs in the world – including, in some cases, acts by other people). Let's call this last category of things *outcomes*. The concept of moral responsibility applies quite broadly; in particular, we hold agents morally responsible both for their own acts and for outcomes. Causal responsibility, on the other hand, applies most fundamentally, if not exclusively, to outcomes. Thus, given that our topic is the relation between causation and moral responsibility, my main focus will be outcomes. The question I will discuss is what, if any, is the connection between moral responsibility and causation as they apply to outcomes. This having been said, as we will see shortly, an account of the conditions under which an agent is morally responsible for an outcome is likely to make reference to

the agent's moral responsibility for his own acts. For an agent's acts are the means by which the agent interacts with the world. The question of when an agent is morally responsible for an act is importantly connected to the question of when an act is free, which I cannot take up here.

## 2. Entailment-Dependence

Imagine that someone is killed in a remote part of the world: someone with whom you have never had contact of any sort. Do we consider you to be morally responsible for his death? Obviously not. Why not? Intuitively, because you weren't causally related to the death in any way: there is no way to link anything you did to the person's death. So here's a very natural idea about the relation between moral responsibility for outcomes and causation: moral responsibility for an outcome *requires* causing it. Moreover, when an agent is morally responsible for an outcome, the existence of causal responsibility partly *explains* the existence of moral responsibility: the agent is morally responsible, in part, *because* he is causally responsible. Given that, on this view, the attributions of moral responsibility both require and depend on the attributions of causal responsibility, we may call this view 'Entailment-Dependence'. Entailment-Dependence is a conjunction of two claims: an *entailment claim* (moral responsibility entails causal responsibility) and a *grounding claim* (the existence of causal responsibility partly explains the existence of moral responsibility).

Now, presumably, to say that an agent caused an outcome is to say that one of the agent's actions or omissions caused the outcome. For example, I may have caused an explosion because I started the fire that caused it or because I failed to put out the fire with my fire extinguisher.<sup>1</sup> Hence, Entailment-Dependence can be fully spelled out as follows:

**Entailment-Dependence:** An agent A is morally responsible for an outcome O only if A is causally responsible for O, i.e. only if one of A's actions or omissions caused O; moreover, the fact that one of A's actions or omissions caused O (partly) explains the fact that A is morally responsible for O.

Entailment-Dependence is the most widely held view about the relationship between moral responsibility for outcomes and causation: it is the view that most people hold, either explicitly or implicitly.<sup>2</sup> As we have seen, it is quite natural and well motivated.

If true, Entailment-Dependence provides part of the foundations for a theory of moral responsibility for outcomes. Given that, on this view, moral responsibility depends on causal responsibility, the question arises whether there is anything else on which it depends, and, if so, what. Whereas it is quite clear that causal responsibility is not sufficient for moral responsibility, it is surprisingly hard to say exactly what else is

needed. In what follows I survey the most salient issues and I point at some potential complications and questions that come up in this connection.

First, moral responsibility for an outcome seems to require, besides causal responsibility for the outcome, moral responsibility for the act leading to the outcome (that act in virtue of which the agent is causally responsible for the outcome). For imagine that someone forces me to take a drug that then makes me set up a bomb in a crowded place. Even if my act of setting up the bomb causes harm, and thus I am causally responsible for the harm, I am clearly not morally responsible for the harm. I am not morally responsible for the harm because I am not morally responsible for the act issuing in the harm (presumably, because my act is not free).

Also, agents can be causally responsible for products of their (free) acts that they couldn't possibly have foreseen would follow from them. Imagine that, as a result of a long and unpredictable series of events, my eating an apple issues in someone's setting up a bomb and causing much destruction. Again, although I am causally responsible for the destruction (I am one of the links in the causal chain leading to it), I am not morally responsible for it. Presumably, then, we should also require that the agent could foresee (or, alternatively, should have foreseen) that the outcome would probably ensue from his acting, or failing to act, in the relevant way.

So moral responsibility for an outcome requires: causal responsibility for the outcome (if Entailment-Dependence is true), moral responsibility for the act leading to the outcome, and, finally, the agent's capacity to foresee that the act was likely to lead to an outcome of such type. But the combination of all these factors is not yet sufficient to guarantee moral responsibility for the outcome. For, as with almost any other causal account of a philosophically interesting concept, there seems to be a problem of 'causal deviance'. Imagine a drunk driver who causally contributes to an accident in the following way: an edgy horse who was walking along the road gets nervous because of the mere presence of the car and throws his rider onto the ground, injuring him. The horse would have panicked even if the driver of the car had not been drunk. In fact, he would have panicked even if someone else had been driving the car. In this case, the drunk driver's driving is a cause of the harm, he is morally responsible – in fact, to blame – for driving (given that it was done under the influence), and, moreover, he could foresee or should have foreseen that his driving would likely result in the occurrence of harm. However, he is not morally responsible for the harm, for the respect in which his act (his driving) was faulty was causally irrelevant to the production of the harm (Feinberg 427). Hence, we could try to amend the account in the following way: not only could the agent foresee (or should have foreseen) that harm would probably follow, but he could foresee (or should have foreseen) that it would follow in (roughly) the way it did follow. How to make this more precise, however (or even whether this is the right way to go), is not totally clear.<sup>3</sup>

Even further revisions might be required.<sup>4</sup> The preceding discussion shows that giving an account of moral responsibility for outcomes in terms of causal responsibility is no easy task. However, if Entailment-Dependence is true, then at least we have the beginnings of such an approach: an important ingredient of moral responsibility for outcomes is causal responsibility. So let us return now to this basic starting point of the standard view: the idea that moral responsibility requires causal responsibility. Is this simple thought right?

### 3. *Quasi-Entailment-Dependence*

An important thing to notice is that, for the simple thought to be right, it must be possible for the omissions of agents to be causally efficacious. Otherwise, agents wouldn't ever be morally responsible for outcomes by omission, whereas they clearly are (as when a lifeguard intentionally fails to jump into the water to save a drowning child). Now, on some views of causation, omissions (and absences in general) cannot be causes or effects. According to such views, the only things that can be causes and effects are events (positive occurrences with definite spatio-temporal locations), whereas omissions are not events but absences thereof.<sup>5</sup> If omissions cannot be causes, then moral responsibility doesn't always require causal responsibility: when an agent is morally responsible for an outcome by omission, he is not causally responsible for it.

Now, how plausible is the view that omissions aren't causally efficacious? And what should we take the link between causal and moral responsibility to be, if omissions aren't causally efficacious? Let us start with the latter question.

Imagine that I promised to water a friend's plant and failed to do so; as a result, the plant died and I am morally responsible for its death. On the view under consideration, we cannot explain my moral responsibility on the grounds that (among other things) I caused the plant's death, since we are assuming that my failure to water it didn't cause its death. But we can say the following. The actual scenario contains a causal process leading to the plant's death (a biological process consisting in the plant's gradually withering until it dies). Although my failure to water the plant wasn't part of that causal chain, in the closest possible world where I do water the plant, my watering the plant cuts such process off (and, as a result, the plant doesn't die). In other words, although I didn't cause the plant's death, I failed to do something that is such that, had I done it, it would have causally interfered with the process that in fact led to the plant's death. Thus we may understand claims about causation by omission, instead of as genuine causal claims, as *counterfactual* claims about causation, or as 'quasi-causal' claims.<sup>6</sup> Correspondingly, we shouldn't think that moral responsibility requires and depends on causation; instead, we should think it requires and depends on *either* causation *or* quasi-causation. On

this view, the reason I am morally responsible for the plant's death is, among other things, that I failed to do something that *would* have interfered with the actual process leading to the plant's death. (Again, other things are required to account for my moral responsibility. Presumably, I must be responsible for not interfering with such a process, I must have foreseen or should have foreseen that by not interfering with it the plant would die, etc.)

Let's call this view 'Quasi-Entailment-Dependence'. The view says:

**Quasi-Entailment-Dependence:** A is morally responsible for O only if A is either causally or quasi-causally responsible for O, i.e. only if one of A's actions or omissions either caused or quasi-caused O; moreover, the fact that one of A's actions or omissions either caused or quasi-caused O (partly) explains the fact that A is morally responsible for O.

Note that, just like Entailment-Dependence, Quasi-Entailment-Dependence regards the notion of causal responsibility as more basic than that of moral responsibility, and attempts to account for the latter in terms of the former. The difference is that whereas Entailment-Dependence accounts for moral responsibility in terms of *actual* causal connections, Quasi-Entailment-Dependence accounts for it in terms of *actual or counterfactual* causal connections.<sup>7</sup>

Now, is Quasi-Entailment-Dependence plausible? Common sense dictates that omissions can be causes, e.g., a doctor's failure to treat a patient can cause the patient's death. Correspondingly, common sense dictates that an agent's moral responsibility is partly determined by his causal responsibility regardless of whether it's by omission or commission. For example, we think that the shooter who inflicted bullet wounds on the patient's body and the doctor that fails to treat the patient once he's at the hospital are both responsible for the patient's death (although maybe to different degrees), and they are both responsible because they both causally contributed to his death (although in different ways). In light of this, we would need strong reasons to give up the view that moral responsibility requires causal responsibility, even in cases of omission.

Helen Beebe has argued that the intuitive support for the idea that omissions can be causes is, despite appearances, not very strong. For not only does common sense dictate that certain omissions are causes; it also dictates that certain other omissions *aren't*. However, Beebe argues, no respectable theory of causation can successfully capture both intuitions. For instance, imagine that the gardener promised to water my plant but forgot, and then the plant died. Beebe claims that any reasonable theory that entails that the gardener's failure to water my plant caused its death would also entail that the *Queen of England's* failure to water it caused its death. After all, if the gardener's failure caused it, it must be because, *had he watered the plant, it would have lived*. However, this is also true of the

Queen of England: had she watered the plant, it would have lived.<sup>8</sup> This shows, for Beebee, that regardless of whether we say that omissions are causes, we'll have to depart from common sense: either by saying that omissions are never causes, or by saying that they are causes much more often than we would have thought.

However, the departure from common sense that the view that omissions cannot be causes requires might be more significant than Beebee realizes. As Jonathan Schaffer has argued, if omissions (and absences in general) are never causes, then many other actions and events won't be causes either ('Causation by Disconnection'). For in many cases an action or event only causes an outcome *via* an intermediate absence (in fact, Schaffer argues that this is what happens in some paradigmatic cases of causation, like shootings: a bullet's piercing a heart causes a death by causing the absence of blood in the brain). If this is right, then we are probably better off accepting the claim that omissions can be causes and then distinguishing the gardener's responsibility from the Queen of England's responsibility on non-causal grounds.

#### 4. Entailment-Reverse-Dependence

Now, there might be another alternative. As Beebee pointed out, the reason we have trouble distinguishing the gardener's causal responsibility from the Queen of England's causal responsibility is that, whereas it is true of the gardener that, had he watered the plant, the plant wouldn't have died, the same is true of the Queen of England (had she watered it, it wouldn't have died). Thus the attempt to explain the gardener's causal responsibility in counterfactual terms entails that the Queen of England was also causally responsible. But there might be another way of grounding the gardener's causal responsibility, one that doesn't similarly entail that the Queen of England was causally responsible.

Judith Thomson has suggested the following principle as an alternative way of grounding the gardener's causal responsibility:

(P) If x is at fault for y, then x causes y.

For Thomson, fault is a normative notion but it is not an essentially moral notion: artifacts and other physical objects can be at fault for certain outcomes and thus cause them, according to (P). Hence, the gardener causes the plant's death but, also, an alarm clock that fails to go off causes me to be late to a meeting, and a gland that fails to secrete an enzyme to protect against a certain disease causes the disease. For Thomson, all of these claims are grounded in the fact that something was *defective* in some way. Note that there is nothing defective about the Queen of England's failure to water my plant: she wasn't at fault in failing to water my plant, so, whereas (P) entails that the gardener caused my plant's death, it doesn't similarly entail that the Queen of England caused it.

Sarah McGrath has made a proposal along similar lines. Her proposal is, roughly, the following. The gardener's failure caused the plant's death because he was *supposed to* water the plant, thus preventing its death. More precisely, had the gardener watered the plant, his watering it would have prevented the plant's death, and in a *normal* way (in McGrath's words, the gardener's watering the plant is a 'normal would-be preventer' of the plant's death; McGrath 141). Similarly for the alarm clock and the gland: the alarm clock's failure to go off caused my being late because the alarm clock's going off is a normal would-be preventer of my being late, and the gland's failure to secrete the enzyme caused the disease because its going off is a normal would-be preventer of my being late. Again, this proposal avoids the result that the Queen of England caused my plant's death: she wasn't supposed to water my plant in order to prevent its death (her watering it is not a *normal* would-be preventer of the plant's death).<sup>9</sup>

Note that, on Thomson and McGrath's proposals, in contrast with the previous proposals (Entailment-Dependence and Quasi-Entailment-Dependence), we don't ground the gardener's moral responsibility for the plant's death in his causal responsibility for it. On the contrary: we ground his causal responsibility in his moral responsibility. More precisely, we ground his causal responsibility in his instantiating a certain *normative* property with respect to the outcome of the plant's death (the property of being at fault for it, or of being such that acting otherwise would have resulted in the normal prevention of the outcome), and we claim that being morally responsible is one of the ways in which something can instantiate such a normative property.

Here, then, is a third proposal about the link between causation and moral responsibility:

**Entailment-Reverse-Dependence:** A is morally responsible for O only if A is causally responsible for O; moreover, the fact that A is morally responsible for O explains the fact that A is causally responsible for O (for being morally responsible is a way of instantiating normative property N, and A's instantiating N explains A's causal responsibility).<sup>10</sup>

To clarify: the view isn't that the explanation of the fact that an agent is causally responsible for an outcome is *always* that he is morally responsible for it. This would be too strong: as we have noted, it is easy to be causally responsible for something without being morally responsible for it. Instead, we should take Entailment-Reverse-Dependence to be the view that, *when* (if) an agent is morally responsible for an outcome, the agent is thereby also causally responsible and his moral responsibility explains his causal responsibility.

Note that, like the standard view (Entailment-Dependence), Entailment-Reverse-Dependence claims that moral responsibility entails causal responsibility; the difference is in what grounds what: whereas for the standard view causal responsibility grounds moral responsibility, for Entailment-

Reverse-Dependence it's the other way around. In other words, Entailment-Reverse-Dependence agrees with the standard view about the entailment claim but disagrees about the grounding claim. (Quasi-Entailment-Dependence, on the other hand, agrees with the grounding claim but disagrees with the entailment claim.) The following table summarizes the three views with respect to the entailment claim and the grounding claim:

	<b>Entailment</b>	<b>Grounding</b>
Entailment-Dependence	Yes	Yes
Quasi-Entailment-Dependence	No	Yes
Entailment-Reverse-Dependence	Yes	No

As I have defined it, Entailment-Reverse-Dependence is only a schema that doesn't acquire full content until we specify what N is. As we have seen, Thomson and McGrath made their own suggestions as to how to understand N. But it is not clear that their proposals will do the required work, as they stand. Notably, a problem with both proposals is that they don't seem to be able to handle cases of *commendable* (as opposed to blamable) omissions. Imagine that I am running for mayor, there is another candidate with good chances of winning and I come across some information about the other candidate's personal life that would lead some prejudiced people to vote against him. I am tempted to propagate this information but I refrain. As a result, his image remains unscathed and he wins the election. We would like to say that my failure to propagate the information was a cause of his image remaining unscathed, and thus of his winning the election. However, I wasn't at fault, I wasn't supposed to act otherwise, and my propagating the information is not a 'normal' would-be preventer of his image remaining unscathed, or of his winning the election. Thus both Thomson and McGrath fail to explain why I am causally responsible for the outcome in this case.<sup>11</sup> In order for Entailment-Reverse-Dependence to get off the ground, then, we should probably look for a more encompassing normative property to do the job.

Now, is the basic idea behind Entailment-Reverse-Dependence plausible? As Thomson herself acknowledges, it seems to 'get things backwards', in the sense that, assuming it is true that moral responsibility entails causation, it is not because causation rests on moral responsibility but because, on the contrary, moral responsibility rests on causation (Thomson 102). In other words, Entailment-Reverse-Dependence seems to reverse the order of explanation. Granted, we are sometimes led to make causal judgments by our moral judgments, as when the causal contribution of an agent is made very salient to us by virtue of the fact that we want to commend what he did, or censure it. However, this doesn't make the causal *facts* dependent on the moral *facts*. Indeed, the natural view is clearly the opposite: the moral facts are dependent on the causal facts.

To illustrate, consider a scenario where someone is morally responsible (say, to blame) for an outcome. Suppose causal responsibility is necessary (although not sufficient for) moral responsibility; thus our agent is also causally responsible. How should we account for his causal responsibility? It seems that there are two options. We could try to account for it in terms of his moral responsibility. Or we could try to account for it in independent terms and then account for his moral responsibility in terms of his causal responsibility.<sup>12</sup> To use an example by Fischer and Ravizza (*Responsibility and Control* 1), imagine that you come home one evening and discover that a valuable vase has been broken. Compare two cases: in the first case, you discover that a houseguest has purposefully broken the vase; in the second case, you discover that your cat has broken it. Whereas the houseguest is morally responsible, the cat isn't; however, they are both (in their respective scenarios) causally responsible. Now, how should we account for the guest's causal responsibility? Should we account for it in terms of his moral responsibility? It seems not: it seems that we should account for it in roughly the same way that we would account for the *cat's* causal responsibility in the other scenario. For instance, if we account for the cat's causal responsibility by tracing a chain of events linking the cat's behavior to the outcome, then it seems that we should account for the guest's causal responsibility in the same way. After all, if the cat is causally responsible because something he did can be linked, by means of a chain of events, to the vase's shattering, then the same should be true of the guest. Now, if this is the right way of proceeding in this case, then we should expect it to be the right way of proceeding in every case: in general, we should proceed by giving an independent account of causal responsibility, and we should then account for moral responsibility in terms of causal responsibility.

Although Thomson sees the problem that her view generates, she says she is prepared to bite the bullet because she sees no other way of accounting for the causal contribution of agents in cases like the gardener case. One could try to explain the gardener's causal contribution counterfactually (by claiming that the reason the gardener is a cause is that, had he watered the plant, the plant wouldn't have died) but, as we have seen, this has the unwanted consequence that the Queen of England was also causally responsible. Thomson seems to think that this consequence would be too high a price to pay. However, one wonders if it is really that high, compared to the price one has to pay by embracing a view like the one she proposes. After all, the belief that causal facts are not grounded in moral facts is a very fundamental belief in our conceptual scheme.<sup>13</sup>

Let us review the three views that we have discussed so far. First, we discussed the standard view, Entailment-Dependence, which takes moral responsibility to entail, and to be grounded in, causal responsibility. Then we looked at Quasi-Entailment-Dependence, which agrees with the standard view about the grounding claim but presents a challenge to the

entailment claim. The challenge is based on the idea that omissions cannot genuinely cause anything, and thus it is possible to be morally responsible for an outcome without being causally responsible for it (in cases of moral responsibility by omission). However, as we have seen, the claim that omissions (and absences in general) are not causally efficacious can, and probably should, be resisted. Another way of challenging the standard view is to claim that the grounding claim should be reversed: moral responsibility doesn't depend on causal responsibility but, on the contrary, causal responsibility depends on moral responsibility. We saw how this view can be motivated by appeal to, again, cases of omission. This gave rise to Entailment-Reverse-Dependence. However, the claim that causal responsibility is grounded in moral responsibility is a hard bullet to bite.

This might suggest that the standard view is the best view to have, all things considered. However, as we will see, there are other challenges to it. In particular, there are other challenges to the entailment claim (which don't rely on the claim that omissions are never causally efficacious). As a result, a fourth view of the relation between causal and moral responsibility might be needed. I will conclude by providing a sketch of such a view.

### 5. *Transmission-Dependence*

A way of challenging the entailment claim without objecting to the causal efficacy of omissions is to claim that moral responsibility comes apart from causation in cases of 'causal overdetermination'. A paradigmatic case of overdetermination is this: Gunman1 and Gunman2 shoot at Victim and their bullets reach him at the same time, killing him. In these circumstances, some people think, neither is a cause of the death (e.g., Bennett). However, we would want to hold each of them morally responsible for it. Why do some people think that neither is a cause in this case? Because, if we were to say of either that he is a cause, then we would have to say of both that they are, but then we would be left with an overabundance of causes. For each of the gunmen's acts is independently sufficient to explain why the effect came about.

In response, however, it could be said that committing ourselves to more than one sufficient cause is, at least in some cases, not problematic. Moreover, it doesn't seem implausible to say, in this particular case, that each of the gunmen was a cause of Victim's death.<sup>14</sup>

But there might be other cases for which the claim that neither agent is a cause is much more plausible (and where someone is still morally responsible). Consider, for instance, the following scenario (discussed in Hart and Honore 218–25). Two people, Badman1 and Badman2 want Victim dead. Victim is about to go on a trip to the desert and he plans to leave with his canteen filled with water. When Badman1 finds out about Victim's plans, he replaces the water with poison. Then, ignoring

what Badman1 did, Badman2 punches a hole in Victim's canteen and drains the liquid (the poison, which he takes to be water) out of it. Once in the desert, Victim dies of thirst. Here is an argument to the effect that this is a case of moral responsibility without causation. *Someone* is morally responsible for Victim's death: either Badman1 or Badman2, or both (if it hadn't been for 'the human hand', Victim would still be alive). However, neither Badman1 nor Badman2 is a cause. Badman1 isn't a cause because the poison he put in the canteen was causally inefficacious, since Victim never got to drink it (given what Badman2 did). And Badman2 isn't a cause because the hole he punched in the canteen was causally inefficacious (given what Badman1 did): the hole could only have been causally efficacious if it had drained water out of the canteen, but, since Badman1 had replaced the water with poison, it never got to do that. In other words, Badman1 and Badman2's actions rendered each other causally inefficacious.<sup>15</sup>

Here is yet another case that might be seen as a counterexample to the entailment claim (for a more extensive discussion, see Sartorio, 'Disjunctive Causes'). Imagine that there is an out-of-control train going along a track. Victim is trapped in the track far ahead. The train approaches a switch. If the switch is flipped, the train will turn onto a side track and run on it for a while. However, the side track loops around and reconverges with the main track. Unfortunately, the reconvergence occurs before the place where Victim is standing, so there is no way to save Victim's life. If the switch is flipped, does the switching cause the death? Under the circumstances, many people think not (see, e.g., Paul; Yablo; Sartorio, 'Causes as Difference-Makers'). However, had the side track been the only way for the train to reach Victim (e.g., if part of the main track had been disconnected right after the switch), then the switching would have been a cause. In other words, it seems that the switching is a cause if, and only if, there is no alternative route to the outcome. Now imagine that the main track had been disconnected for repairs and was reconnected this morning by an evil man, Badman1, who realizes that sometimes things go badly, people get trapped in tracks, and get hit by trains, so he reconnected the track hoping that an accident of this sort would occur. When, later on, Victim finds himself trapped in the track and the runaway train reaches the switch, a second evil man, Badman2, flips the switch, thinking that the main track is still disconnected and thus the side track is the only way for the train to reach Victim. The train runs on the side track for a while, then back on the main track, and ends up killing Victim. Again, here is an argument that this is a case of moral responsibility without causation. *Someone* (either Badman1 or Badman2 or both) is morally responsible for Victim's death. If it hadn't been for the malicious interventions of the two evil men, Victim wouldn't have died: the train would have derailed while running on the main track, and Victim would have been spared. However, neither Badman1 nor Badman2 is a cause.

Badman1 isn't a cause because his reconnecting the main track was causally inefficacious (given what Badman2 did: since Badman2 redirected the train, the train never got to run on that part of the main track). And Badman2 isn't a cause because his redirecting the train was causally inefficacious (given what Badman1 did: since Badman1 reconnected the main track, the train would have reached Victim all the same).

Other cases could be offered as counterexamples to the entailment claim.<sup>16</sup> All of these cases suggest that moral responsibility might not require causal responsibility. If so, assuming that we still wish to retain the grounding claim (the idea that moral responsibility somehow rests on causation), we should rethink the idea that the way the grounding works is by virtue of causal responsibility being one of the necessary conditions that are jointly sufficient for moral responsibility. If the grounding claim is true, it must be true in some other way.

In light of the above examples, what could the relationship between causation and moral responsibility amount to? It will help to look at some features that the examples have in common. Here are some common features. First, in all these cases we want to hold some moral agent(s) responsible for the outcome because, intuitively, the outcome was a product of human intervention (not, say, an accident of nature). Second, two moral agents are involved, but neither is causally responsible for the outcome because the agents' individual contributions 'cancel each other out': the reason one agent isn't a cause is that the other agent acted, and vice versa. Still, it is clear that the agents bring about the outcome 'together', in particular, the outcome wouldn't have happened had it not been for the fact that they acted in the ways they did. So, how can we hold the agents (or at least one of them) morally responsible, if the outcome was not an upshot of the agents' individual causal contributions? We can do so by claiming that it was still an upshot of the agents' *collective* contribution. As long as we can make the agents (or at least one of them) morally responsible for that collective behavior, we can hold the agents (or at least one of them) morally responsible for the outcome.

On this view, causation is not one among many individually necessary but jointly sufficient conditions for moral responsibility. Instead, causation is the 'vehicle of transmission' of moral responsibility: if the agent is morally responsible for an outcome's cause (and if other conditions are met),<sup>17</sup> the agent can be said to be morally responsible for the outcome. Crucially, the outcome's cause for which the agent must be responsible needn't be one of the agent's own actions/omissions; it can be, instead, a larger collective behavior involving the behaviors of different individuals.

Note that, on this view, the agents' moral responsibility is still explained in causal terms (although it is not necessarily explained in terms of the agent's individual causal powers). In other words, the view still attempts to ground attributions of moral responsibility in attributions of causal responsibility, but without claiming that an agent's causal responsibility for

an outcome is one of the necessary conditions that are jointly sufficient for his moral responsibility for the outcome.

The new view can be labeled ‘Transmission-Dependence’. Briefly, it claims:

**Transmission-Dependence:** A is morally responsible for O only if A is morally responsible for a cause of O; moreover, the fact that A is morally responsible for a cause of O (partly) explains the fact that A is morally responsible for O.

It is important to note that, just like the other views we have examined here, Transmission-Dependence is a view of moral responsibility for *outcomes* (events and states of affairs ‘out there in the world’). This avoids an infinite regress, in which we would otherwise fall: the view doesn’t say that, in order for an agent to be morally responsible for an outcome, he would have to be morally responsible for a cause of the outcome, and for a cause of that cause, and so on *ad infinitum*. The view doesn’t apply to causes of the outcome that aren’t themselves outcomes. A clear example of something that is not an outcome is, again, an action or omission by the agent. But collective acts or behaviors (those that involve the actions and omissions of the agent and the actions and omissions of other agents) also aren’t outcomes. Hence, in the examples given above, assuming that we can hold the agent(s) morally responsible for the relevant collective behavior on some other grounds, we can explain his (their) moral responsibility for the final outcome by appeal to Transmission-Dependence.

There are several questions that arise for this view. Unfortunately, I don’t have the space to try to answer them all here. So, in what follows, I briefly review some of those questions and I sketch the general way in which I think they should be answered (see Sartorio, ‘How to Be Responsible’ for further discussion).

First, how should we determine what the relevant ‘collective behavior’ is, in a case where an agent is morally responsible for an outcome but not causally responsible for it? Is the relevant collective behavior a ‘conjunctive’ behavior, or a ‘disjunctive’ behavior, or something of a different kind? In ‘Disjunctive Causes’ I argue that we should conceive of it as a disjunctive behavior (which supports the somewhat surprising view that causes can be disjunctive). For example, in the train example described above, the cause of Victim’s death is the fact that either Badman1 reconnected the main track or Badman2 redirected the train to the side track (a disjunctive fact).

Second, in virtue of what are the agents morally responsible for such collective behavior? Even if, as I have pointed out, we shouldn’t expect Transmission-Dependence to give us an answer to this question, the following worry remains. Imagine that we say that they are morally responsible because what each of them did was ‘part of’ that collective behavior, as it seems natural to suggest. Given that the collective behavior caused

the outcome, wouldn't this imply that each of their acts was part of the 'total cause' of the outcome, and thus, that each of their acts was a cause of the outcome? But, if so, each agent would be causally responsible for the outcome, and we're back with Entailment-Dependence. In response, I think that we should not explain the agents' moral responsibility for the collective behavior by saying that what each of them did was *part of* the collective behavior. Instead, we should say that the agents are morally responsible for the collective behavior (understood as a disjunctive behavior) because the collective behavior is *logically entailed* by each of the individual behaviors. Presumably, if Badman1 is morally responsible for his act of reconnecting the main track, then he can also be morally responsible for a logical consequence of his act, such as the fact that *someone* reconnected the main track, and he can be morally responsible for it in virtue of the fact that it logically follows from his act. Similarly, if Badman1 is morally responsible for reconnecting the main track, he can also be morally responsible for the fact that either he reconnected the main track or Badman2 redirected the train, which logically follows from Badman1's act, and he can be morally responsible for this disjunctive behavior in virtue of the fact that it logically follows from his individual behavior. In other words, the proposal would be to say, roughly, that agents are morally responsible for outcomes in the way specified by Transmission-Dependence, and that they are morally responsible for the causes of those outcomes, either because they are acts that they perform freely, or because they are logical consequences of those free acts.

At any rate, if we are persuaded by the counterexamples to the entailment claim presented above, Transmission-Dependence suggests itself as a plausible alternative to the standard view about the relation between causation and moral responsibility. It avoids the counterexamples by rejecting the entailment claim, but at the same time it preserves the grounding claim, which is very intuitively plausible.

### *Acknowledgment*

For valuable discussion and suggestions, thanks to Juan Comesaña, Gabriel Uzquiano, and a referee for *Philosophy Compass*.

### *Short Biography*

Carolina Sartorio is Assistant Professor of Philosophy at the University of Wisconsin, Madison. Her research lies at the intersection of metaphysics and ethics. She has published articles on causation, agency, and moral responsibility in the journals *Noûs*, *Philosophical Perspectives*, *The Journal of Philosophy*, and *Philosophical Studies*. She holds a B.A. in Philosophy from the University of Buenos Aires and a Ph.D. in Philosophy from the Massachusetts Institute of Technology.

## Notes

\* Correspondence address: Philosophy Department, University of Wisconsin at Madison, 5195 Helen C. White Hall, 600 North Park Street, Madison, WI 53706, USA. Email: [acsartorio@wisc.edu](mailto:acsartorio@wisc.edu).

<sup>1</sup> I am setting aside the possibility of irreducible 'agent-causation' (a form of causation that, some philosophers argue, is not to be understood in terms of event-causation; see Chisholm).

<sup>2</sup> For example, Feinberg develops this type of view in detail in his article. And Wolf endorses this view when she writes: 'when we hold an individual morally responsible for some event, we are doing *more* than identifying her particularly crucial role in the causal series that brings about the event in question. We are regarding her as a fit subject for credit or discredit *on the basis of* the role she plays' (166; emphases added).

<sup>3</sup> Feinberg notes that there is some controversy about the necessity of the non-deviance condition; in his work he defends its inclusion from potential objections.

<sup>4</sup> For example, an interesting complication arises if, as some philosophers have suggested, there is an asymmetry between the conditions giving rise to praise and those giving rise to blame (see, e.g., McCann; McCann offers a philosophical defense of an asymmetry exposed by Knobe). According to such philosophers, whereas intentionally bringing about a bad outcome might be sufficient for blame, intentionally bringing about a good outcome is not sufficient for praise. If so, our theory of moral responsibility might need to provide different conditions for bad and good outcomes. Also, importantly, whether further revisions are needed might depend on one's view of causation. Fischer and Ravizza identify the following problem for a similar kind of approach (suggested, but not endorsed, by Heinaman): suppose Elizabeth has launched a missile towards a city; Joan cannot stop it but she can deflect it to a less populous area of the city. Fischer and Ravizza argue that, if she does this, she is not morally responsible for the fact that the missile hit the city (somewhere or another); however, she is responsible for her act of deflection, her act is causally sufficient for the outcome, and she could foresee that the outcome would come about (moreover, one could add: the causal chain leading to the outcome is not deviant). However, there are views of causation on which Joan's act is not a cause of the outcome (the missile's hitting the city somewhere or another). For discussion of these cases in the context of the connection between causation and moral responsibility, see Sartorio, 'Causes as Difference-Makers'.

<sup>5</sup> See, e.g., Beebe; Dowe, 'Counterfactual Theory'.

<sup>6</sup> This is Phil Dowe's terminology in his 'Counterfactual Theory'. This is a simplified version of Dowe's account. For a more exhaustive discussion, see also *Physical Causation*.

<sup>7</sup> Dowe writes: 'Quasi-causation may also, in my opinion, help track moral responsibility in just the way that causation does' ('Counterfactual Theory' 225).

<sup>8</sup> The idea seems to be that whether one promised to water a plant cannot plausibly matter to whether one caused its death by failing to water it (more on this later).

<sup>9</sup> Although there is much in common between Thomson and McGrath's proposals, there are also important differences between them. One difference is that, whereas McGrath regards the fact that an event of a certain type is a normal would-be preventer as both sufficient and necessary for the omission of an event of that type to be a cause, Thomson's principle (P) only sets a sufficient condition on causation. Another difference is that in some cases McGrath's concept of the abnormal seems to come apart from Thomson's concept of being at fault. For example, McGrath seems to think that any disruption of a statistical regularity is abnormal in the relevant sense, but Thomson would probably not say that there is always something defective or at fault when this happens.

<sup>10</sup> Why 'explains' instead of 'partly explains'? Because, in this case, no other conditions seem to be needed: if the moral claim grounds the causal claim, and if the moral claim entails the causal claim, then the moral claim seems sufficient to explain the causal claim. Another example of someone who endorses a view of this type is Alvarez.

<sup>11</sup> This is more obviously a problem for McGrath than for Thomson given that, as I pointed out (see n.9), Thomson, unlike McGrath, only intends to offer sufficient conditions for causation. But it is also a problem for Thomson. For Thomson's principle (P) provides an explanation in

the gardener case but not here; however, it seems that what we want is a unifying explanation of both phenomena.

Could McGrath say that my propagating the information is a normal would-be preventer of the other candidate's winning the election (maybe on the grounds that it is 'expected' for people – or politicians – to do this)? I am not sure. Maybe there is a sense in which it is 'normal' for politicians to maliciously spread information to win elections. But it seems doubtful that the sense in which this is normal is the same sense in which it is normal for, say, gardeners to water the plants they have agreed to water.

<sup>12</sup> A third option would be to take causation as a primitive. I will ignore this option here.

<sup>13</sup> In addition, we could try to 'explain away' the intuition that the Queen of England isn't a cause of the plant's death in pragmatic terms (the reason we don't see her as a cause of the plant's death is that she is a highly non-salient cause; by contrast, the gardener is a salient cause, given that we were expecting him to water the plant).

<sup>14</sup> Even Kim, who is famous for his 'principle of explanatory exclusion', thinks that a commitment to more than one sufficient cause might be justified in cases of causal overdetermination (see Kim). For an argument that there is nothing problematic with overdetermination, or, in fact, with widespread overdetermination, see Schaffer ('Overdetermining Causes').

<sup>15</sup> I am assuming that Victim would have died in pretty much the same way (and at around the same time) if he had drunk the poison. If this isn't a plausible assumption, the example would have to be changed accordingly.

<sup>16</sup> See Sartorio ('How to Be Responsible') for a counterexample involving omissions (one that assumes that omissions can be causally efficacious).

<sup>17</sup> What conditions? Presumably, similar conditions to those that came up in our discussion of the standard view (Entailment-Dependence).

### Works Cited

- Alvarez, M. 'Agents, Actions, and Reasons'. *Philosophical Books* 46.1 (2005): 45–58.
- Beebe, H. 'Causing and Nothingness'. *Causation and Counterfactuals*. Eds. J. Collins, N. Hall, and L. A. Paul. Cambridge, MA: MIT Press, 2004. 291–308.
- Bennett, J. *Events and their Names*. Indianapolis, IN: Hackett Publishing Company, 1988.
- Chisholm, R. 'Human Freedom and the Self' (1964). Lindley Lecture, University of Kansas. Reprinted in *Free Will*. Ed. G. Watson. Oxford: Oxford UP, 1982.
- Dowe, P. 'A Counterfactual Theory of Prevention and "Causation" by Omission'. *Australasian Journal of Philosophy* 79.2 (2001): 216–26.
- . *Physical Causation*. New York, NY: Cambridge UP, 2000.
- Feinberg, J. 'Sua Culpa'. *Doing and Deserving: Essays in the Theory of Responsibility*. Princeton, NJ: Princeton UP, 1970. 187–221.
- Fischer, J. M. and M. Ravizza. *Responsibility and Control*. Cambridge: Cambridge UP, 1998.
- . 'Responsibility for Consequences'. In *Harm's Way: Essays in Honor of Joel Feinberg*. Eds. J. Coleman and A. Buchanan. Cambridge: Cambridge UP, 1994. 183–208.
- Hart, H. L. A. and A. M. Honore. *Causation in the Law*. Oxford: Oxford UP, 1959.
- Heinaman, R. 'Incompatibilism without the Principle of Alternative Possibilities'. *Australasian Journal of Philosophy* 64 (1986): 266–76.
- Kim, J. 'Mechanism, Purpose, and Explanatory Exclusion'. *Philosophical Perspectives* 3 (1989): 77–108.
- Knobe, J. 'Intentional Action and Side Effects in Ordinary Language'. *Analysis* 63 (2003): 190–3.
- McCann, H. 'Intentional Action and Intending: Recent Empirical Studies'. *Philosophical Psychology* 18 (2005): 737–48.
- McGrath, S. 'Causation by Omission: A Dilemma'. *Philosophical Studies* 123 (2005): 125–48.
- Paul, L. 'Aspect Causation'. *Journal of Philosophy* 97.4 (2000): 235–56.
- Sartorio, C. 'Causes as Difference-Makers'. *Philosophical Studies* 123 (2005): 71–96.
- . 'Disjunctive Causes'. *Journal of Philosophy* 103.10 (2006): 521–38.
- . 'How to Be Responsible for Something Without Causing It'. *Philosophical Perspectives* 18 (2004): 315–36.

- Schaffer, J. 'Causation by Disconnection'. *Philosophy of Science* 67 (2000): 285–300.
- . 'Overdetermining Causes'. *Philosophical Studies* 114.1/2 (2003): 23–45.
- Thomson, J. 'Causation: Omissions'. *Philosophy and Phenomenological Research* 66.1 (2003): 81–103.
- Wolf, S. 'The Real Self View'. *Freedom within Reason*. Oxford: Oxford UP, 1990. 23–45.  
Reprinted in *Perspectives in Moral Responsibility*. Eds. J. M. Fischer and M. Ravizza. Ithaca, NY: Cornell UP, 1993. 151–69.
- Yablo, S. 'De Facto Dependence'. *Journal of Philosophy* 99 (2002): 130–48.