

Indeterministic Compatibilism*

Carolina Sartorio

(University of Arizona)

Abstract: Although free will compatibilists are typically focused on arguing that determinism is compatible with free will, most compatibilists also think that indeterminism is compatible with free will too—and I am one of those compatibilists. In this paper, I'll look at this issue from the perspective of a compatibilist view I've defended elsewhere (*Causation and Free Will*, Oxford University Press 2016): a view that takes our freedom to be a function of the actual causal histories of our behavior. In the first part of the paper I argue that, assuming this view, it follows that indeterminism is in fact compatible with free will. Still, the assumption of indeterminism gives rise to some novel and interesting questions concerning the nature of indeterministic causation. The second part of the paper is concerned with motivating and discussing those questions.

Keywords: Indeterminism, compatibilism, free will, indeterministic causation

1. Introduction

Free will compatibilists are typically focused on arguing that the truth of determinism would not undermine our freedom and responsibility. Most compatibilists also think, however, that the truth of determinism is not *required* for free will—in other words, they think that the truth of indeterminism is compatible with free will too. Unsurprisingly, given compatibilism's main

aim, little work has been done on this aspect of compatibilism.¹ Still, it is important to think about this, if one is interested in developing a view of free will that doesn't hinge on determinism being actually true or false—and I am one of those compatibilists.

In this paper, I will look at this issue from the perspective of a compatibilist view that I have developed and defended elsewhere (Sartorio 2016): a view where the type of freedom or control required by responsibility is accounted for in terms of responsiveness to reasons, and where responsiveness to reasons is in turn a feature that is directly reflected in the causal histories of our behavior. Thus, this is a view according to which acting freely is a matter of our acts having the right kinds of causes, and it is a form of compatibilism because the right kinds of causes can be deterministic. Now, under the assumption that our acts fail to be determined, the causal histories of our acts will be, at least partly, indeterministic. An examination of the compatibility of this view with indeterminism will then lead us into a discussion of how our free will could be grounded in causes that are not fully sufficient for their effects, as well as into the intriguing nature of indeterministic causation.

In the first part of the paper I will argue that, assuming this compatibilist view of free will, indeterminism does not constitute an obstacle to our freedom and responsibility. What is important, on this view, is the existence of causal histories that are rich or robust enough to ground our reasons-responsiveness, but not in a sense that requires them to be deterministic. Still, as we will see, the assumption of indeterminism gives rise to some novel and interesting questions. The second part of the paper will be concerned with motivating and discussing those questions.

One issue that I cannot take up in this paper is the question of whether there is a (potentially new and worrisome) problem of luck that may arise for compatibilists as a result of the application to indeterministic contexts. This is an important issue that I cannot get into here, since it would require its own extended treatment. My main focus will instead be on the more basic or fundamental issue of how compatibilist views could be applied to indeterministic contexts, and the special questions that arise at that earlier stage.

2. The compatibility of compatibilism with indeterminism

I will illustrate with an example of a kind analyzed in detail by Kane (1996, chapter 8), an incompatibilist (and libertarian) about free will, one that involves “self-forming” or will-setting acts. Imagine that at a certain point in your life you are forced to choose between satisfying your own self-interested goals and providing needed assistance to someone else. Imagine, for example, that you are on your way to a very important meeting, one that is likely to advance your career in significant ways, when you see a wounded man who needs your immediate assistance. If you stop to help him, you won’t make it to your meeting on time, and you will miss the only chance you have (and will likely have in years) to advance your career in the way you wanted and you think you deserve. On the other hand, if you don’t stop to help the wounded man, you have reason to believe that, although he will survive, his wounds might get infected, which could result in complications for his long-term health. So, this is a case where you have strong reasons to do the selfish thing (continue on to your meeting) and also strong reasons to do the selfless thing (help the man).² We are to imagine that, given those compelling reasons pulling in opposite directions, you feel very torn about what to do. But, regardless of

what you end up deciding to do, this is an important decision that will help shape your future character in significant ways (by turning you into a more selfish or selfless person, say); hence the label “self-forming.” Moreover, we are to imagine that your decision is causally undetermined, so it is in fact compatible with the past and the laws of nature that you will make either decision.

Call this case **Choice**. Kane argued that, if there are instances during our life stories where we make important decisions of this kind, then those self-forming choices can constitute the locus of our free will (Kane 1996). But, Kane argued, it is crucial that these be causally undetermined decisions, since this is the only way in which we can be the ultimate originators or ultimate sources of our wills, which he thinks is a fundamental requirement for acting freely. Thus, according to Kane, your decision in **Choice** can be made freely to the extent that you make that decision voluntarily and rationally, or on the basis of compelling reasons, and to the extent that it fails to be causally determined.

The details of Kane’s view are not important for our purposes here. The reason I will focus on a case like **Choice** is that, despite being an indeterministic case, by design it has the potential for meeting the conditions for free action set out by the compatibilist view that I am assuming here. And this is regardless of what you end up deciding to do. If you decide to stop to help the man, then you do it for reasons, and those reasons are arguably part of the causal history of your act, even if that causal history is indeterministic. Similarly, if you decide not to stop to help the man, you also do it for reasons, and those reasons, again, are arguably part of the causal history of your act, even if the causal history is indeterministic. In fact, examples like **Choice** are of special interest because they suggest that there can be instances where, although

it is genuinely undetermined *what* we will do (say, whether we will do A or B instead), we act freely regardless of what we do (if we do A or if we do B)—roughly, because we are acting for reasons either way. (This is also Kane’s own position on this, although he would add that the fact that your choice was undetermined was, in addition, a requirement for you to have free will.)³

Now, things are in fact more complicated than this because being reasons-responsive in the sense required to act freely and to be responsible is not simply a matter of acting for the *actual* reasons that you had. This is too simplistic, for it would entail that someone who acts for reasons automatically acts freely, which is clearly false. For example, a compulsive behavior can be done for reasons (say, to satisfy an irresistible urge) but it is not free. Compulsive behaviors are not free because they are not reasons-responsive in the relevant sense.

At this point, different reasons-responsiveness views give different accounts of how compulsive and other unfree behaviors come apart from free behaviors. But the common strategy used by reasons-responsiveness views is, roughly, to expand the set or pattern of reasons to which one has to be sensitive in order to be sufficiently reasons-responsive. On the view that I favor, one has to look at the role played by, in addition to the actual reasons, the *absence* of various other (counterfactual) reasons. In a nutshell, and simplifying quite a bit, the view can be stated as follows:

Causal Reasons-Responsive Compatibilism (CRRC): Reasons-responsiveness is causal sensitivity to an appropriate range of reasons and absences of reasons, one that

includes *actual* reasons to do what you are doing as well as the *absence of* (counterfactual) sufficient reasons to refrain from doing what you are doing.⁴

Very roughly, the idea is this. Imagine that you decide not to help the wounded man in **Choice** on the basis of the selfish reasons. The thought is that, if you do this freely (not, for example, as a result of some irresistible compulsion to act selfishly), then in making that choice you are also responding or being causally sensitive to the absence of a range of sufficient reasons to do otherwise (reasons to help the man). Imagine, for example, that you would have stopped to help the man if you had reason to believe that others were watching (you care very much about what others think about you). Or imagine that you would have stopped to help the man if you had been informed of the existence of a substantial financial reward for doing so, one that would help you and your family immensely. Etcetera. In fact, conditions like these didn't obtain. But the point is that we can account for the fact that you acted freely (and not, for example, compulsively) by thinking about the role played by the absence of reasons of this kind in an explanation of your behavior. Given that you were not acting compulsively and you were disposed to act differently if conditions of that kind had been present, part of the explanation of your actual behavior seems to be that such conditions did not in fact obtain. This suggests that free behaviors are behaviors whose causal histories are quite rich in that they include, in addition to the actual reasons, the absence of several other (counterfactual) reasons. And it is in this way that free behaviors differ from unfree behaviors such as things that we may do compulsively.⁵

Although CRRC is a compatibilist view of free will, and this means that the causal histories in question can be deterministic, they don't in fact need to be deterministic. According to CRRC, the difference between free and non-free behaviors amounts to a difference in the content of the causal histories of those behaviors; it does *not* amount to a difference in the type of causal relation that ties those contents (the causes of the behaviors) to the behaviors. For example, even if the past and the laws didn't guarantee that you would make the selfish choice in **Choice**, if you do end up making that choice and your choice is indeterministically caused by an appropriate range of actual reasons and absences of counterfactual reasons, your choice is free, according to CRRC.

The easiest way to see that free will is compatible with indeterminism according to CRRC is by focusing on a simple version of **Choice**, which I will call **Choice 1**, and to which I already alluded above (I discuss other more complex variants later in the paper). Imagine, again, that you decide not to help the man on the basis of the selfish reasons, and that this choice was undetermined by the past and the laws. Also, imagine that you are psychologically constituted in such a way that, if others had been watching or if there had been a substantial financial reward for helping the man, then you *would* have chosen to help him. (That is, I am assuming that this is a *deterministic* relation: if either of those conditions had obtained, then the chance that you'd decide to help the man in those circumstances would have been 1.) In that case it seems clear that the absence of reasons of that kind is part of what accounts for your choice to not help the man in the actual scenario—that is, those absences of reasons are part of the causal history of your choice. As a result, CRRC entails that you were responding to reasons in the relevant sense when you made the choice and, thus, that you acted freely.

But notice that all of this is consistent with the initial assumption that the causal history of your choice was indeterministic. For the past and the laws didn't have to guarantee that you would choose to do the selfish thing in the actual conditions, where those other reasons were not present (we may still assume that the past and the laws were compatible with your making the opposite choice). Thus, CRRC entails that you acted freely in **Choice 1**, even if the causal history of your choice was indeterministic.

We can imagine a similar variant of the case where you make the opposite choice (you choose to help the man). Imagine, for example, that you wouldn't have made that choice (i.e. the chance of your making that choice would have been null) if you had reason to believe that your spouse would divorce you, or that your whole family will be ruined as a result of your missing the only chance you had to advance your career. Then the causal history of your choice would include the absence of facts like these. As a result, you would be reasons-responsive, according to CRRC. And, again, this is so even if the causal history of your choice was indeterministic—even if those reasons and absences of reasons, in conjunction with any other causes of the choice, did not guarantee that you would act selflessly.

This strongly supports the compatibility of CRRC with indeterminism. If acting freely is a matter of having the right kind of actual causal history, one including the relevant combination of reasons and absences of reasons, then acting freely is in fact consistent with the causal histories of our acts not being deterministic.

Now, all of this only works under certain assumptions about the nature of causation—and, in particular, indeterministic causation—which I have been implicitly taking to be true. More discussion of these assumptions would be helpful. Also, it's not clear what would happen

if one tried to generalize to other cases that have a more complex structure than **Choice 1**. I take these issues up, in turn, in the following two sections.

3. Indeterministic causation and probability-raising

One assumption I have obviously been relying on is that causation *can* be indeterministic: causes needn't be sufficient for their effects (even when we take the "whole cause" of an effect, or the combination of all the factors that causally contributed to it). This assumption is widely accepted nowadays.⁶ And this isn't something that we can only conceive happening at the microscopic or quantum level, where the possibility of indeterminism being real usually comes up. If a terrorist manages to build an indeterministic bomb (one that has a chance smaller than 1 of going off) and the bomb actually goes off, the terrorist causes the explosion, even if the explosion wasn't causally necessitated by the terrorist's act in conjunction with anything else.⁷ Surely, events like this still have causes, even if they are not sufficient causes. Plus, it is easy to see how indeterministic causal relations of this kind could potentially ground the responsibility of agents. For example, if the terrorist is morally responsible for the explosion, his responsibility would be partly grounded in the fact that he caused the explosion to happen, even if he didn't guarantee that it would happen.

But how are we to make sense of indeterministic causation? A natural and quite popular way to think about it is in terms of objective probabilities (or chances) and, in particular, in terms of the idea of *probability-raising*.⁸ The main motivation behind this thought is that, even if indeterministic causes don't guarantee the occurrence of their effects, they can still make them more likely to occur (than if they had been absent). In light of this, they can make a

contribution to the occurrence of those effects, by virtue of having raised the probability of their occurrence, *when* and *if* they occur. (Of course, if the effects don't come about, as it's bound to happen in some cases, then there is no such causation simply because causation is factive and the effects did not in fact occur.)

Understanding indeterministic causation in terms of probability-raising can help us see, for example, how causes can bring about their effects even in cases where the effects had only a small chance of occurring. Imagine that it was much more probable, given the past and the laws, that you would help the man in **Choice** than that you would not. Imagine, for example, that this time the chances were 0.9 and 0.1, respectively, but you still decided to do the unlikely thing, the selfish thing, on the basis of the same set of reasons (and thus, we would like to say, freely).⁹ An account in terms of probability-raising can explain how those reasons caused your choice even if the event of your making that choice was highly unlikely. For example, the selfish reasons may have raised the chance of your making the selfish choice from 0 to 0.1, if the choice had no chance at all of occurring in the absence of those reasons (because you wouldn't have at all been motivated not to help the man if you didn't have a very important meeting to attend, one with potentially life-changing implications). In that case it is clear that the reasons caused the choice.

Now, accounts of indeterministic causation in terms of probability-raising face important challenges. This is not the place to review them all.¹⁰ For now I will just touch on one of them that is relevant for my purposes here. It is the problem posed by certain kinds of *preemption* cases (Lewis 1986: 179). Imagine that Suzy, an unreliable terrorist, and Billy, a reliable terrorist, are simultaneously trying to make a bomb go off. They each do this by

throwing a switch that is part of a mechanism that is hooked up to the bomb; however, whereas Suzy's mechanism is very unreliable (it only has a 0.1 chance of success), Billy's mechanism is very reliable (it has a 0.9 chance of success). Imagine that Suzy's and Billy's mechanisms are also connected with each other in such a way that, when Suzy throws her switch, it shuts off Billy's reliable mechanism at the same time that it starts its own unreliable process. Imagine that, despite this, the unlikely happens and the bomb still goes off. This case threatens to undermine the idea that indeterministic causation can be understood in terms of probability-raising. For notice that Suzy's throwing her switch doesn't raise, but instead considerably lowers, the probability of the bomb going off; however, it still clearly causes that outcome. The unreliable process started by Suzy *preempts* the reliable process started by Billy, and thus it (and not Billy's process) causes the explosion, despite having made the explosion less probable.

Causation theorists have tried to deal with this problem in different ways. One main strategy has been to tinker with the probability-raising requirement in a way that unreliable preempting causes like Suzy end up being probability-raisers, in the relevant sense. One way to do this is to understand the probability-raising requirement as claiming that, when there is more than one potential causal route or path to an outcome, in assessing whether an event raises the probability of the outcome, one must hold fixed facts involving the other potential paths. In other words, the relevant sense of probability-raising is in an important way path-specific, or restricted to a particular causal pathway.¹¹ This way of understanding the probability-raising requirement yields the desired result in the preemption case. For, holding fixed the fact that Billy's process is no longer active after a certain time, Suzy's act of throwing

the switch does raise the probability of the outcome (because Suzy's process is the only active process that could potentially lead to it). But not everybody would agree that this type of strategy fully addresses the problem, and this (as other challenges that arise for the probability-raising view) is still an issue of much debate.

The reason this is relevant for our purposes here is that we can easily imagine a kind of preemptive structure underlying cases like **Choice**, one that mimics other cases widely discussed in the free will literature ("Frankfurt-style" cases, originally from Frankfurt 1969). Imagine, again, that the chance that you would help the man is much higher than the chance that you would do the selfish thing, say, 0.9 versus 0.1. Now add a preempted alternative involving a resourceful and evil neuroscientist. Imagine that the neuroscientist wanted you to do the selfish thing, so earlier in the day he installed a chip in your brain that started a process that deterministically guarantees that you will do the selfish thing, by causing you to make the selfish choice, but only if the (unreliable) indeterministic process started by your own selfish reasons doesn't do it on its own.¹² Imagine that the unlikely happens, and you decide to do the selfish thing on your own, on the basis of the selfish reasons. So, again, the unreliable process preempts the reliable process (which in this case is a completely reliable or fully deterministic process). Your selfish reasons caused your choice even if they didn't raise the probability of that choice. In fact, in this case, the chance that you would make the choice was already 1 by the time you considered those reasons, so the reasons clearly couldn't have raised that chance any further. Call this case **Frankfurt-style Choice**.

The CRRC account of free will would say that all that matters to your freedom is the actual causes of your behavior. So, if your choice is caused by your own deliberation and selfish

reasons, and not by the process started by the neuroscientist, then you act freely in **Frankfurt-style Choice** (and this is despite the fact that you couldn't have done otherwise). Moreover, views like CRRC are typically *motivated* by intuitions about Frankfurt-style examples. For those examples are taken to show that all that matters to freedom is actual causal histories, or actual explanations of behavior, and not something like having alternative possibilities of action or being able to do otherwise. Thus, the thought that an agent in a Frankfurt-style case is responsible for his choice because he made the choice for his own reasons, or because his own reasons caused the choice, is central to a view like CRRC. If this causal claim couldn't be sufficiently supported, then this would be a serious blow to the view. But, as we have just seen, it seems that we cannot substantiate such a causal claim by appealing to a probability-raising view of indeterministic causation, which is the most natural way to try to understand that form of causation.

However, this isn't a problem for the CRRC view. For, as we have also seen, this is a problem that probability-raising views have with preemption cases in general. And it is a problem for those views precisely because it seems clear what the causal structure of those cases is, and the probability-raising view has trouble accommodating it. The apparent failure of those views does not make us doubt the causal structure of such cases; if anything, it's the other way around: the causal structure of the cases makes us doubt the truth of those views (or to look for refinements). In this respect, the preemption problem that arises for probability-raising views of indeterministic causation is just like the preemption problem that arises for most reductive views of *deterministic* causation (which is, in and of itself, a big problem).¹³ In

both cases, they are problems because the views don't seem able to accommodate the causal facts, which we take to be clear (or clear enough).

To conclude this section, the main assumption that underlies the extension of the compatibilistic view to indeterministic settings is just this: there can be indeterministic causation, and, in particular, there can be indeterministic causal processes that have the potential to ground our reasons-responsiveness. In this section I argued that this assumption is not undermined by worries concerning the prospects of a probability-raising account of causation, or, in general, by any uncertainty concerning the underlying metaphysics.

4. Indeterminism, causal underdetermination, and causal indeterminacy

Still, indeterminism raises some interesting new questions. To establish the compatibility point all we needed was the simplest case of a certain kind, **Choice 1**. Recall that a feature of that case was that, although the causal history of the choice was in fact indeterministic, there were other (counterfactual) relations that I assumed to be deterministic. In particular, I was assuming that, had certain sufficient reasons to do otherwise been present, then you would have done otherwise (it was deterministically guaranteed that you would). You actually chose to do the selfish thing, and this was not determined, but, had other people been watching or had there been a substantial financial reward for helping the man, then you would have helped the man (the chance of your helping the man would have been 1). The reason I focused on that case is that, under those assumptions, it seems perfectly clear that the absence of those reasons partly accounts for your actual choice, and thus it is easy to see that you are reasons-responsive, according to a view like CRRC.

But, of course, this isn't the most natural case to consider under the assumption of indeterminism. What if those chances would not have been 1, but some number in between 0 and 1? That is to say, what if there is some chance, but not a full guarantee, that you would have responded to those reasons if they had been present? Are you reasons-responsive in that case? Call this new version of the case **Choice 2**.

CRRC would say that you are reasons-responsive in **Choice 2** to the extent that those absences of reasons still caused your choice. But, *did* those absences of reasons cause your choice in this case? Here matters are much less clear than before. Imagine, for example, that the chance of your choosing to do the selfish thing on the basis of the selfish reasons was 0.5, and so was the chance of choosing to help the man on the basis of the selfless reasons. Imagine that you actually chose to do the selfish thing. Moreover, imagine that, had other people been watching, then the chance of your helping the man would have been higher (say, 0.7 instead of 0.5) and your chance of doing the selfish thing lower (0.3 instead of 0.5). In these conditions, did the absence of that reason actually cause your choice? Did the fact that nobody was watching causally contribute to your choosing to do the selfish thing?

Note that a probability-raising account of indeterministic causation would entail that such a causal connection exists. For the fact that nobody was watching raised the probability that you would choose to do the selfish thing (it was more likely that you would do the selfish thing given that nobody was watching than if someone had been watching). Plus, you in fact did the selfish thing.

However, there is reason to be wary here. For one, as we have seen, we shouldn't blindly trust everything that a probabilistic account says. But, also, people's judgments are likely

to be less clear at this point. Some would agree with the judgment entailed by the probabilistic account. But others would disagree. Instead, they would suggest that there are *two* distinct possibilities that are compatible with the setup of the case: one where there is a causal connection of the kind we were imagining and one where there is not. And the probabilities don't settle which possibility is the actual one.

This is sometimes called the problem of *underdetermination* (for reductive theories of causation), and it has been discussed in the causation literature as part of an argument for primitivism about causation. Basically, the thought is that probability-raising examples can be used to show that the non-causal facts underdetermine the causal facts. The causal facts are primitive in that they are not reducible to other facts—in particular, they are not reducible to probabilities or probability-raising. All we can say in underdetermined cases is that it is possible that a certain causal relation exists and also possible that it doesn't, or how likely it is that there is one and how likely it is that there isn't one.

Consider, as another example to motivate this, the following scenario discussed by Schaffer: Two sorcerers, Merlin and Morgana, simultaneously cast spells with a 0.5 chance of turning the prince into a frog. Each spell has an independent 0.5 chance of succeeding. If the prince turns into a frog, who caused it to turn into a frog? Call this case **M&M**.¹⁴

In magic-involving cases like **M&M** it is assumed that, when spells work, they work directly (not through any intermediate events). This is important because it means that we cannot hope to establish which spell was the cause by looking for causal intermediaries or their absence; *all* we have is the probabilities. In light of this, a probability-raising account would simply entail that Merlin and Morgan are both causes because both of them were probability-

raisers. But somebody with primitivist leanings would protest that there are in fact *three* distinct possibilities in this case: either Merlin is a cause, or only Morgana is a cause, or both are causes. And there are no non-causal facts about the situation that determine which is the actual scenario. Again, the causal facts are brute: they are what they are, and they cannot be reduced to non-causal facts (in particular, probabilities).

Now, note that **Choice 2** is relevantly like **M&M** in that all we have to go by is the relevant probabilities. If, for example, the absence of other people watching is indeed a cause of your choice, then this is an instance of absence causation. As a result, there won't be an ordinary process linking cause and effect via intermediate events. When there is absence causation, it works "directly" or without any causal intermediaries.¹⁵ This is why the cases are equally prone to eliciting primitivist intuitions about causation.

In turn, other philosophers with reductivist leanings would counterargue that this results in unacceptable metaphysical burdens. It commits us to causal differences that "float free" in that they are not grounded in any other feature of the world. Moreover, they would argue, the primitivist intuitions elicited by these kinds of cases can be explained away. Lewis (1986: 180-3) tries to explain them away by arguing that they are the result of not taking genuine indeterministic chances seriously, and thus of assuming that there are features of the world that remain hidden from us when we give the probabilistic descriptions of the cases. In turn, Schaffer (2008: 89) argues that we can explain away the primitivist intuitions by attributing them to errors of reification (mistaking mere concepts for real things). Given that we have the concepts needed to describe the different causal possibilities, we tend to think that they are real possibilities when in fact they aren't.

So, what would the reductivist say about the causal structure of these cases? Lewis sticks to his guns and says that whenever an event C raises the chance of another event E that actually occurs, it automatically follows that C is a cause of E (Lewis 1986: 180). In other words, probability-raising is sufficient for causing, at least when the effect actually occurs.¹⁶ It follows that, in **M&M**, both spells caused the enchantment, and, in **Choice 2**, the fact that nobody was watching caused your choice to do the selfish thing.

Schaffer (2008) would agree with Lewis about these specific causal judgments. But, as he notes elsewhere, the view that probability-raising is sufficient for causation needs further refinement (Schaffer 2000). Schaffer asks us to consider the following “overlapping” variant on the **M&M** case, which we may call **Overlapping M&M**: Merlin casts a spell with a 0.5 chance of turning the *king and prince* into frogs. Morgana casts a spell with a (probabilistically independent) 0.5 chance of turning the *prince and queen* into frogs. As it happens, the king and prince, but not the queen, turn into frogs. Now consider the actual event of the *prince* turning into a frog. Schaffer asks: What caused that event? Clearly, it was Merlin’s spell and not Morgana’s. However, Morgana’s spell also raised the probability of that event happening. So, raising the probability is not sufficient for causing an actual event. The reductivist story in terms of probability-raising needs further refinement.

Schaffer also notes that the reductivist story may have to include the possibility of causal indetermination in order to account for the structure of some special cases. Imagine yet another variant on the **M&M** case (taken from Schaffer 2000: n. 21), which we may call **Enhanced M&M**. Simply add to the description of the **M&M** case that, when more than one spell works, the effect is enhanced, say, the prince-turned-into-frog becomes extra-green.

Imagine that this time the enhanced effect does not obtain. So, we know that this isn't an overdetermination case where both spells were causes. In this case the primitivist would say that either Merlin or Morgana caused the effect and it's a brute fact who did. But the reductivist will have to say that it is indeterminate who did: there is simply no fact of the matter.¹⁷

As suggested by the brief discussion above, this is a lively and ongoing controversy. I think it is fair to say that there is no consensus on what the causal facts are in these cases. But note that there is also quite a bit of uncertainty about the *responsibility* facts themselves. Are you, in **Choice 2**, responsible for making the selfish choice?

This is much less obvious than in **Choice 1**. In **Choice 1**, it was perfectly clear that you were reasons-responsive because it was perfectly clear that the relevant absences of reasons causally contributed to your choice (and thus that you were not acting compulsively, for example). But, in **Choice 2**, all of this is less clear. If all we can say is that you *might* have responded to the sufficient reasons to help the man, or that the probability that you would help the man would have been *higher* than it would have been otherwise, without being a full guarantee, is this enough to satisfy the reasons-responsiveness requirement? This doesn't seem obvious one way or the other.

Now, those with reductivist leanings about indeterministic causation should probably say that it is. For, again, given that those absences of reasons considerably raised the probability of your making the selfish choice, and given that you actually made the selfish choice, it follows from the probability-raising account that those absences of reasons caused

your actual choice. After all, this is basically what indeterministic causation *is*, on these kinds of views. If that is the case, then you are reasons-responsive, and thus responsible, in **Choice 2**.

On the other hand, those with primitivist leanings would say that there are two distinct possibilities consistent with the setup of the case: one where the absences of reasons caused the choice and one where they didn't. The former possibility would make you reasons-responsive but the latter would not. And it is a brute fact which of these possibilities is the actual one. Hence, although there is a fact of the matter about which possibility is in fact actualized, we don't know (and in principle cannot know) what it is. As a result, we don't know (and in principle cannot know) if you are indeed responsible in **Choice 2**.

In sum, in this case it is much less clear that you are responsible, and this is because the underlying causal facts are less clear. Even if you are a committed reductivist about causation, and for that reason you believe that you are indeed responsible in **Choice 2**, you may use the primitivist intuitions (which everybody agrees have some force) to explain away the initial uncertainty about this case. Following CRRC, you may say that it is less obvious that you are responsible (although you are in fact responsible) because it is less obvious that the causal history of your choice includes everything that is needed to make you reasons-responsive.

Could there be cases where there is simply no fact of the matter as to whether an agent is responsible? This depends, again, on your views on causation. Imagine a variant on **Choice** that shares the structure of the scenario that drew Schaffer to commit to causal indeterminacy (**Enhanced M&M**). Let this case be **Enhanced Choice**: As in the **Enhanced M&M** case, there are two indeterministic potential causal routes to your making the selfish choice, each with a 0.5 chance of succeeding. One of them involves a spell by Merlin, but the other involves the

relevant absences of reasons to help the man. Imagine that when they are both causally active this results in an enhanced effect (say, you act particularly selfishly in that you leave the scene without even feeling worried about the man's condition). Imagine that this time the enhanced effect doesn't obtain, so we know that this is not an overdetermination case where both potential causes are causally active.

In this case, the uncertainty about your responsibility seems even more pronounced. The primitivist would explain it by noting that there are two distinct possibilities, one where you are responding to reasons and one where you are manipulated by Merlin, and we have no way of telling which is the actual one. A reductivist like Schaffer, on the other hand, would explain it by arguing that this is a case of fundamental causal indeterminacy: there is no fact of the matter as to which was the cause. So, according to one of the views, it would follow that there is a fact of the matter as to whether you are responsible; we just don't know what it is. According to the other view, it would follow that there simply is no fact of the matter. Either way, the uncertainty or indeterminacy about the causal structure grounds the uncertainty or indeterminacy about your responsibility.

5. Conclusions

As we have seen, the assumption of indeterminism raises some interesting questions about the nature of indeterministic causation and its application to our theories of responsibility. But note that, regardless of how those questions are answered, the discussion in the previous section reinforces the idea that responsibility is closely tied to causation (via reasons-responsiveness) in the way posited by CRRC. For it is arguably due to that close connection between responsibility

and causation that we see that, when the causal facts are less clear, so are the responsibility facts. Thus, the remarks in the previous section are further confirmation of the idea that our free will is grounded in the actual causal histories of our behavior, as CRRC says.

In particular, I think this provides further reason to prefer a *causal* version of reasons-responsiveness to a more traditional “counterfactualist” version like Fischer and Ravizza’s (a view that appeals to counterfactual scenarios).¹⁸ Very roughly, Fischer and Ravizza’s view states that the relevant mechanism is reasons-responsive when it issues in the agent’s doing otherwise in *some* counterfactual scenarios where sufficient reasons to do otherwise are present. As far as I can tell, this condition is easily met in all of these scenarios that we have been considering. For example, in **Choice 2**, there are certainly counterfactual scenarios where the reasons to do otherwise are present and your practical reasoning results in your helping the man. So the requirement of reasons-responsiveness seems to be met in this case, and thus the account seems to entail that you are responsible for your selfish choice.¹⁹ But, again, it is not clear that this is the right result. Perhaps it is, at the end of the day, but matters are not as clear as with other simpler cases (such as **Choice 1**). Plus, a counterfactualist account like Fischer and Ravizza’s account doesn’t have the resources to explain the source of that lack of clarity.

Again, I think that the source of this uncertainty is that what we want to know, in order to determine whether you are responsible in **Choice 2**, is whether in the *actual* case you were responding to certain absences of reasons. In other words, it is a matter of actual (causal or explanatory) relevance. Of course, I don’t mean to suggest that counterfactual possibilities are simply useless in answering this kind of question. Typically they are not. But in *this* case they do appear to be useless—or so the primitivist would say. Again, the primitivist intuition (which

others also agree we have, even if they try to explain it away) is that those kinds of facts don't settle whether an actual causal connection obtains. So, by appealing to the uncertainty about the causal connection, a causal account of reasons-responsiveness can accommodate the uncertainty in responsibility about **Choice 2**, whereas a counterfactualist account cannot.

More generally, I think this supports CRRC over other compatibilist accounts of free will. As we have seen, if it is less clear that agents act freely in some of the cases we have reviewed, it is because it is less clear that the relevant *absences of reasons* played a role in accounting for their behavior. Thus, this helps bring out the importance of the role played by absences of reasons in grounding free will, which is unique to CRRC. And note that this is something that we can only see by thinking about indeterministic cases. For this lack of clarity about the causal structure of the situation doesn't arise in deterministic settings; it only arises in indeterministic settings given the probabilistic nature of those cases and the special metaphysical questions they give rise to. Thus, reflecting on the indeterministic case can be particularly illuminating at the time of formulating a promising compatibilist account of free will, and it can help support CRRC over other forms of compatibilism.

* For helpful feedback, thanks to Al Mele, Michael McKenna, the students in our metaphysics seminar at the University of Arizona (co-taught with Michael in the Spring of 2020), and audiences at the University of Maryland, Florida State University, and the College of William and Mary. Thanks, also, to Marco Hausmann and Jörg Noller for inviting me to contribute to this volume.

¹ Mackie 2018 and McKenna Ms. are two notable exceptions. Both Mackie and McKenna are mostly concerned with addressing the problems of luck and control that compatibilists would inherit from libertarians, given the assumption of indeterminism (and, in McKenna's case, given the assumption of determinism too). As I note below, this is an important issue that unfortunately I cannot take up in this paper.

² This is so even if it might be the case that one set of reasons outweighs the other—for example if helping the man were the only morally permissible thing to do in the circumstances. We don't need to decide this issue here.

³ Mackie (2008) and McKenna (Ms.) also focus on a case of this kind to motivate a similar point. Mackie writes about that kind of case: "But if, whichever of these things she does, it will be a decision for which she has (or takes herself to have) good reasons, why shouldn't the *two-way compatibilist* [a compatibilist who thinks free will is compatible both with both determinism and indeterminism] say that, whichever way she decides, she does so freely—at least if certain other standard compatibilist conditions are fulfilled...?" (Mackie 2008: 281). As will be apparent later, I think those "other standard compatibilist conditions" are where the main action lies.

⁴ See Sartorio 2016 (chapter 4) for a full development of the view. A standard—causalist—theory of agency is assumed throughout the book. As I explain in chapter 2, the assumption that absences can be causally efficacious is only a simplifying assumption, one that can be replaced with an assumption involving other explanatory kinds of relations that absences can participate in, if causation isn't one of them. The main idea is that absences of reasons must be part of what *accounts for* or *explains* our behavior when we act freely. The simplest way of understanding this idea is, of course, in terms of causal explanation.

⁵ Notice that the relevant counterfactual reasons must be *sufficient* reasons to do otherwise (or else we shouldn't expect the behavior of a reasons-responsive agent to be explained by the absence of such reasons). But also note that sometimes reasons-responsive agents fail to respond to the *actual* sufficient reasons to do otherwise (**Choice** could be an example of this if, for example, stopping to help the wounded man were in fact morally required in the circumstances). All that is required to be sufficiently reasons-responsive is sensitivity to an appropriate *range* of reasons or absences of reasons, and these needn't include the actual sufficient reasons.

⁶ Anscombe (1971) forcefully argued for this, and many others have claimed to find the main idea extremely plausible.

⁷ Lewis gives an example of this kind in his 1986: 176.

⁸ This can be done in terms of conditional probabilities (the classical example is Suppes 1970) or in terms of counterfactual conditionals with objective chances in the consequent (the classical example is Lewis 1986, Postscript B).

⁹ You may wonder, though, whether low chances make the problem of luck particularly pressing. Again, unfortunately, this is something I cannot take up here. For a discussion of the problem of luck that arises for compatibilists under the assumption of indeterminism and a comparison with the parallel problem for libertarians, see Mackie 2018 and McKenna Ms.

¹⁰ For discussion of the different challenges, see the contributions in Dowe and Noordhof 2004.

¹¹ See, e.g., Dowe 2004 and Hitchcock 2004a.

¹² This example has the causal structure of a case discussed in Mele and Robb 1998. It involves a special variety of preemption known as "trumping" preemption. Trumping is a particularly tricky kind of preemption because the preempted process is never interrupted or cut off (it just fails to be causally efficacious), and this rules out the use of several strategies that are commonly used to deal with preemption, such as the one discussed above in the text.

¹³ For example, but not exclusively, for counterfactual theories of causation. See the discussion in Paul and Hall 2013, chapter 3.

¹⁴ Schaffer 2008: 88. Similar examples are discussed by Armstrong (1983: 133; 2004: 450), Tooley (1987: 199-202), Woodward (1990), and Carroll (1994: 134-41). Armstrong writes: "Suppose that there are two bombardments of an atom, with the same chance of the atom emitting a particle, which the atom duly does. Does there not seem to be an objective question, which of the two bombardments actually did the job?" (Armstrong 2004: 450) In general, the most interesting cases to think about are cases without causal intermediaries, and the atom case could be one of them. As noted below in the text, magic cases like Schaffer's do that simply by stipulation about how magic works, so they are particularly well-suited for these purposes. Ordinary absence causation doesn't involve magic, but it works in the same kind of way, as I argue next.

¹⁵ This is one of the reasons why some think absence causation is spooky and should be rejected. But recall that the possibility of absence causation isn't strictly speaking necessary for the view to work (see n. 4 above). And note that similar questions would arise if, for example, we were to rephrase the point in terms of the explanatory power of absences instead of in terms of their causal efficacy.

¹⁶ As others have noted, this only seems to work for cases without causal intermediaries (see, e.g., the discussion in Schaffer 2000 and Hitchcock 2004b). But, again, recall that the cases under discussion contain no causal intermediaries.

¹⁷ It wouldn't be the first time a causation theorist commits to this. For instance, Bernstein (2016) offers reasons to believe that *deterministic* causation can also be indeterminate.

¹⁸ See Fischer and Ravizza 1998, and also McKenna 2013. In Sartorio 2016 I give other reasons to prefer a causal account of reasons-responsiveness and free will.

¹⁹ The Fischer and Ravizza view also requires a coherent counterfactual pattern of reasons-responsiveness. So perhaps Fischer and Ravizza could argue that, if you respond to the reason in some worlds but fail to respond to it in other worlds where everything else is equal, then it is not clear that this makes for a coherent pattern of reasons-responsiveness, and thus it is not clear that you are responsible. But I don't think that this would be a satisfying response. For it doesn't seem to get to the heart of the uncertainty about responsibility. Intuitively, the source of the uncertainty is something else: it is that we are uncertain, more fundamentally, about whether you are *in fact* responding to reasons.

Reference list

- Anscombe, G. E. M. 1971. *Causality and Determination: An Inaugural Lecture*. London: Cambridge University Press.
- Armstrong, D. 1983. *What is a Law of Nature?* Cambridge: Cambridge University Press.

-
- Armstrong, D. 2004. "Going through the Open Door Again: Counterfactual versus Singularist Theories of Causation." In J. Collins, N. Hall, and L. A. Paul (eds.), *Causation and Counterfactuals* (Cambridge, Mass: The MIT Press), pp. 445-57.
- Bernstein, S. 2016. "Causal and Moral Indeterminacy." *Ratio* 29, 4: 434-47.
- Carroll, J. 1994. *Laws of Nature*. Cambridge: Cambridge University Press.
- Dowe, P. 2004. "Chance-Lowering Causes." In Dowe and Noordhof 2004, pp. 28-38.
- Dowe, P. and P. Noordhof (eds.). 2004. *Cause and Chance: Causation in an Indeterministic World*. London: Routledge.
- Fischer, J. and M. Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge University Press.
- Frankfurt, H. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66, 23: 829-39.
- Hitchcock, C. 2004a. "Routes, Processes, and Chance-Lowering Causes." In Dowe and Noordhof 2004, pp. 138-51.
- Hitchcock, C. 2004b. "Do All and Only Causes Raise the Probabilities of Effects?" In J. Collins, N. Hall, and L. A. Paul (eds.), *Causation and Counterfactuals* (Cambridge, Mass: The MIT Press), pp. 403-17.
- Kane, R. 1996. *The Significance of Free Will*. New York: Oxford University Press.
- Lewis, D. 1986. "Causation." In *Philosophical Papers II*, New York: Oxford University Press, pp. 159-213.
- Mackie, P. 2018. "Compatibilism, Indeterminism, and Chance." *Royal Institute of Philosophy Supplement* 82: 265-87.

-
- McKenna, M. 2013. "Reasons-Responsiveness, Agents and Mechanisms" in D. Shoemaker (ed.) *Oxford Studies in Agency and Responsibility*, vol. 1 (New York: Oxford University Press).
- McKenna, M. Ms. "Facing the Luck Problem for Compatibilists."
- Mele, A. and D. Robb. 1998. "Rescuing Frankfurt-Style Cases." *The Philosophical Review* 107, 1: 97-112.
- Paul, L. A. and N. Hall. 2013. *Causation: A User's Guide*. Oxford: Oxford University Press.
- Sartorio, C. 2016. *Causation and Free Will*. Oxford: Oxford University Press.
- Schaffer, J. 2000. "Overlappings: Probability-raising without Causation." *Australasian Journal of Philosophy*, 78: 40-46.
- Schaffer, J. 2008. "Causation and Laws of Nature: Reductionism." In T. Sider, J. Hawthorne, and D. Zimmerman (eds.), *Contemporary Debates in Metaphysics* (Blackwell Publishing), pp. 82-107.
- Suppes, P. 1970. *A Probabilistic Theory of Causality*. Amsterdam: North-Holland Publishing Company.
- Tooley, M. 1987. *Causation: A Realist Approach*. Oxford: Clarendon Press.
- Woodward, J. 1990. "Supervenience and Singular Causal Statements." In D. Knowles (ed.), *Explanation and Its Limits* (Cambridge: Cambridge University Press), pp. 215-6.