What Time Travel Teaches Us About Moral Responsibility

By Taylor W. Cyr and Neal A. Tognazzini

Forthcoming in *Journal of Ethics and Social Philosophy*; please cite published version.

Abstract: This paper explores what the metaphysics of time travel might teach us about moral responsibility. We take our cue from a recent paper by Yishai Cohen, who argues that if time travel is metaphysically possible, then one of the most influential theories of moral responsibility (i.e., Fischer and Ravizza's) is false. We argue that Cohen's argument is unsound but that Cohen's argument can serve as a lens to bring reasons-responsive theories of moral responsibility into sharper focus, helping us to better understand *actual-sequence* theories of moral responsibility more generally and showing how actual-sequence theorists should respond to a recent criticism.

Philosophers these days tend to favor ecumenical theories. It would be an undesirable feature of a theory of moral responsibility, for example, if it committed its proponents to a consequentialist theory of normative ethics. Likewise, it would be undesirable if a response to the problem of induction committed its proponents to theism. And so on.

The implicit acceptance of this methodological constraint opens up fruitful avenues of research for those inclined to see how a theory in one area of philosophy might have consequences for theorizing in another area. In this paper we'd like to explore one of these avenues. Specifically, taking our cue from a recent paper by Yishai Cohen (2014), we'd like to see what the metaphysics of time travel might be able to teach us about moral responsibility. In his paper, Cohen argues that if time travel is metaphysically possible, then one of the most influential theories of moral responsibility – that of John Martin Fischer and Mark Ravizza (1998) – is false. If Cohen were right, that would be an especially surprising connection between literatures that have for the most part developed independently of each other.[1]

[1] Spencer (2013) and McCormick (2017) are notable exceptions to this generalization.

In what follows we will argue that Cohen is right to think that we can learn something important about moral responsibility from the metaphysics of time travel, but that the true lesson is not quite the one he has in mind. In particular, we will show that although Cohen's argument is unsound, it can nevertheless serve as a lens to bring reasons-responsive theories of moral responsibility into sharper focus, which in turn will help us to better understand *actual-sequence* theories of moral responsibility more generally.

1.

What connects the metaphysics of time travel with theories of moral responsibility are *counterfactuals*. So, let's begin by tracing both topics to their meeting point.

Moral responsibility is often thought to require free will, and free will is often thought to require the ability to do otherwise. Further, the ability to do otherwise is often thought to imply the truth of certain counterfactual claims. Take, for example, the infamous and discredited conditional analysis, according to which someone is able to do otherwise just in case, were they to desire to do otherwise, they would. Here, free will is analyzed in terms of a particular counterfactual.

But even theorists who endorse Harry Frankfurt's (1969) attack on the Principle of Alternative Possibilities – that is, even theorists who deny that moral responsibility requires the ability to do otherwise – still often talk about moral responsibility in terms of counterfactuals. Take, for example, the most detailed and influential theory of moral responsibility on the market: that of John Martin Fischer and Mark Ravizza (1998). Fischer and Ravizza deny that moral responsibility requires the ability to do otherwise; instead, they offer an *actual-sequence* account

of moral responsibility, according to which when an agent is morally responsible, this is wholly in virtue of facts about the way an action is *actually* produced, and not at all in virtue of facts about how things *might* have unfolded or *would* have unfolded in some non-actual possible world. But which actual-sequence facts matter for moral responsibility?

Fischer and Ravizza focus their attention on the so-called *control condition* for moral responsibility (as opposed to, say, the *epistemic* condition, which is also important but not as frequently discussed), and their contention is that an agent has control over what they do just in case their action issues from their *own, moderately reasons-responsive mechanism*. We'll get into some of the details of their account below, but for now it suffices to note that despite their being champions of an actual-sequence account of moral responsibility, Fischer and Ravizza still rely heavily on counterfactuals in spelling out the notion of reasons-responsiveness. Instead of focusing on what the *agent* would do under certain counterfactual circumstances, however, they focus on the reasons-sensitivity of the agent's *decision-making mechanism*, where that mechanism is sensitive to reasons just in case certain counterfactuals hold. This is a subtle argumentative strategy, and it's of course not without its share of critics, but again, we'll save some of the details for later. For now, the point is that theorizing about moral responsibility seems to lead inevitably to a careful consideration of certain counterfactuals.[2]

The same can be said for the metaphysics of time travel. Here the connection is even easier to see, since philosophical discussions about time travel have tended to center around the Grandfather Paradox and other similar worries about the possibility of backward time travel. Briefly, the worry is that if backward time travel is possible, then contradictions could be true. The rough idea is as follows: if backward time travel is possible, then I could travel back in time to

[2] There is an important exception to this claim that we discuss in section 6 below.

visit my grandfather when he was a child, and in that moment, it would be true *both* that I could kill him – what would stop me? – and also that I *couldn't* kill him – since if he had died in that moment, my mother would never have been born, and then I would never have existed, so I wouldn't be there trying to decide what to do in the first place. The fact that I am there in his childhood means he didn't die in that moment, so it looks like no matter how hard I try to kill him, I'll inevitably fail, despite the fact that it I've got everything I would need in order to pull it off.

This is a rough-and-ready presentation of the paradox, so let's not put too much weight on it.[3] The relevant point is that a proper articulation and evaluation of the Grandfather Paradox will require a deep dive into counterfactual reasoning. For example, the scenario sketched above seems problematic in part because it seems to be describing a situation in which the following counterfactual is true: if I were to kill my grandfather, then I wouldn't have existed. That by itself seems to cause trouble for the supposition that I *can* kill my grandfather while I'm time traveling, but we can cause even more trouble for that supposition by endorsing the following principle, inspired by Kadri Vihvelin (1996: 318): S is able to do A only if, had S tried to do A, S would or at least might have succeeded. The funny thing about me and my grandfather is that, *no matter how hard I were to try*, I would fail to kill him. And if the principle just mentioned is correct, then it follows that I *can't* kill him.

One of the perplexing things about backward time travel – at least, cases of it that involve the time traveler visiting their past self or their direct ancestors – is that it makes counterfactuals go all screwy. All of a sudden my own existence appears to hinge on (i.e., counterfactually depend on) the most mundane of events. Parricide is not mundane, of course, but that's just a particularly vivid example. In the *Back to the Future* film franchise, the same basic paradox is explored without

[3] See Wasserman (2018, chapters 3 and 4) for a comprehensive discussion of this and related paradoxes.

parricide, and instead with the simple and accidental event of keeping one's own parents from ever falling in love. But whatever the details of the story, in cases of backward time travel, our usual method for evaluating counterfactual statements seems to lead us into trouble, since facts about the future (that is, about the time traveler's personal past, before they got into the time machine) seem like they must be held fixed – and, to put it simply, we just aren't used to doing that. It's the *past* that's fixed, while the future is *open*. But in cases of time travel, as David Lewis (1976: 151) puts it, facts about the future "masquerade" as facts about the past.

So far, we have explained how our two topics – moral responsibility and time travel – both require careful thinking about counterfactuals, but this falls short of the task we set ourselves in this section, which is to show how the topic of counterfactuals *connects* theorizing about moral responsibility with the metaphysics of time travel. Now that we have the backstory, we can make relatively quick work of that task.

Here's the bottom line: the most influential theory of moral responsibility understands the crucial notion of *control* in terms of the holding of certain counterfactuals that provide the details about whether (and to what extent) an agent's action-producing mechanism is sensitive to reasons, but in cases of backward time travel, counterfactuals that we ordinarily take to be boringly true turn out to be bewilderingly false (or else we have no idea what to say about them). What that means is that there will be time travel stories that will seem, at least at first glance, to provide counterexamples to this theory of moral responsibility. As we have seen, in cases of time travel we can get counterfactuals about the behavior of agents to come out false, *seemingly without interfering with the intrinsic capacities of the agents in question*, and instead just by placing them in the right external circumstances. So, if your preferred theory of moral responsibility is both (a) committed to the truth of certain counterfactuals and (b) ostensibly concerned solely with an

agent's intrinsic psychological capacities, then you probably can't have both of those things at the same time.

In the next section we'll look closely at a detailed version of this worry, raised recently by Yishai Cohen against Fischer & Ravizza's theory of moral responsibility. Our contention will be that although Cohen's argument is unsound, taking it seriously will teach us something important about theories of moral responsibility more generally, especially ones that claim to focus exclusively on the *actual sequence*.

2.

In a recent paper, Yishai Cohen claims if we add one seemingly harmless thesis to the theory of moral responsibility championed by Fischer & Ravizza, then that theory is inconsistent with the metaphysical possibility of time travel. This would be a very odd result to say the least, but it would also be an unattractive result, especially to Fischer & Ravizza, who are explicitly concerned to construct a theory of responsibility that doesn't hinge on "the arcane ruminations" (Fischer 2006: 5) of theoretical physicists (or, presumably those of metaphysicians, either). Moreover, there is fairly wide consensus among contemporary metaphysicians that the usual objections to the metaphysical possibility of time travel fail, so it would be a mark against Fischer & Ravizza's theory if it required them to take a dissenting view.[4] Fortunately, Cohen's attempt to saddle Fischer & Ravizza with this result is unsuccessful. But before we explain why, let's take a closer look at Cohen's argument.

[4] As Cohen (2014: 6, n. 19) notes, Dowe (2000) defends the metaphysical possibility of time travel, and Artzenius and Maudlin (2002) discuss its nomological possibility. For the classic defense of the metaphysical possibility of time travel, see Lewis (1976). For a more recent (and the first book-length) defense of the metaphysical possibility of time travel, see Wasserman (2018).

To see how Cohen's argument works, we need to explain the Fischer & Ravizza account of moral responsibility in a bit more detail. We have already said that Fischer & Ravizza offer an account of the control condition on moral responsibility, and they lay out a set of necessary and sufficient conditions for an agent's exercising that sort of control. They call it *guidance control*, and their account runs as follows:

> An agent exercises *guidance control* over an action just in case the action issues from the agent's own, moderately reasons-responsive mechanism, where a mechanism is moderately reason-responsive just in case it is regularly receptive to reasons and at least weakly reactive to reasons.[5]

The notions of *regular receptivity* and *weak reactivity* here are spelled out in terms of how the mechanism would respond in various counterfactual circumstances:

> A mechanism is *regularly receptive to reasons* just in case there are possible scenarios in which (1) there is sufficient reason to do otherwise, the same kind of mechanism is operative, and the agent recognizes that reason, and (2) the possible scenarios described in (1) constitute an understandable pattern of reasons-recognition.
>
> A mechanism is *weakly reactive to reasons* just in case it is regularly receptive to reasons and, in at least one of the possible scenarios described in the account of regular receptivity, the agent chooses and does otherwise for the reason in question.[6]

These formulations are adequate, but they are also a bit abstract. Here's the basic idea: when a morally responsible agent acts, the process leading up to their action (the "mechanism") is capable of "seeing" the relevant reasons and is also capable of reacting appropriately to those reasons. To figure out whether a mechanism has the relevant capabilities, we look to facts about nearby worlds. So long as there is an intelligible range of possible circumstances in which this particular decision-making process *does* "see" the reasons there are, then we can say that the *actual* decision-making process is *capable* of "seeing" those reasons. Likewise, so long as there is at least one possible

[5] This is our paraphrase of the account elaborated and defended in Fischer & Ravizza 1998. We are setting aside the ownership component of guidance control, since this does not play a role in Cohen's argument, but see Fischer & Ravizza (1998, chapter 8) for their account of ownership.
[6] Again, we are paraphrasing. For the full details, see Fischer & Ravizza (1998: 69-76).

circumstance in which, having "seen" the reasons, the relevant decision-making process kicks into gear and issues in a choice on the basis of those reasons, then we can say that the *actual* decision-making process is *capable* of reacting to those reasons. (The rationale for why receptivity requires an "understandable pattern" whereas reactivity only requires "at least one" relevant possible scenario need not detain us here.)

One of Fischer & Ravizza's key innovations is to distinguish between the *agent* and the *mechanism* by which the agent acts.[7] They do this for two related reasons: (1) they are persuaded by so-called Frankfurt-style counterexamples that an agent can be morally responsible for what they have done even if the agent wasn't able to have done otherwise, and (2) they want to defend a positive theory of moral responsibility that focuses on the capacity to respond to reasons. Since the notion of capacity is a paradigm modal notion, Fischer & Ravizza need to find a way to get modality into their account without giving up on the insight of Frankfurt-style counterexamples. They do this by distinguishing between agents and mechanisms: the agent may not be able to do otherwise, but that doesn't mean the mechanism on which the agent acts isn't *capable* of responding to the relevant reasons.

But it is this very desire to accommodate a modal notion like *capacity* that, according to Yishai Cohen, puts the Fischer & Ravizza account of moral responsibility on a collision course with the metaphysical possibility of time travel. This is because, as we have seen, cases of backward time travel make trouble for our ordinary ways of thinking about counterfactuals. The essence of Cohen's objection is this: we can easily construct a backward time travel story according to which the time traveler seems for all the world to be morally responsible for what they have done, but, due to the metaphysical peculiarities involved in attempting to kill one's

[7] See Fischer & Ravizza (1998: 38-41).

younger self, there does not exist the range of worlds that Fischer & Ravizza say is needed for the agent to be acting on a moderately reasons-responsive mechanism. Here is a modified version of the story that Cohen tells (2014: 3):

> Zoe lives in a peculiar world. First, time travel is nomologically possible. Second, individuals can commit murder merely by *willing that someone die*. However, there is one line of defense available to the would-be victims: they can continue to live simply by willing to nullify the attempted murder. Now, suppose that Zoe travels 20 years into the past to visit a younger version of herself, and suppose that she wills that her younger self die. However, her attempted murder does not succeed, because her younger self wills to nullify the attempt.

Now, Cohen claims that if we think carefully about the relationship that Zoe has to her younger self, we'll see that the mechanism that the younger Zoe acts on cannot be moderately reasons-responsive. It's crucial that these are two person-stages of the very same individual, because that means that the very existence of Zoe-the-time-traveler depends counterfactually on her failure to kill her younger self. With that in mind, we can see that, once Zoe has become a time traveler, there are no worlds in which younger Zoe dies, and hence no worlds in which she refrains from willing to nullify her older self's attempted murder.[8] But if there are no worlds in which she refrains, then *a fortiori* there isn't an "understandable pattern" of worlds in which she sees the reasons to refrain and then acts on them. But it is precisely this pattern of worlds that Fischer & Ravizza say is required for younger Zoe to be morally responsible for her behavior.

The argument is not yet complete, however. All that follows, so far, is that if Fischer & Ravizza are right about moral responsibility, then younger Zoe is not morally responsible for willing to nullify her older self's attempted murder. For this story to constitute a *worry* for Fischer & Ravizza, we need some independent reason to think that their view gives us the *wrong* verdict

[8] This is a bit too quick, actually, since there may be worlds in which young Zoe is killed by her future self but is then somehow resurrected. (Thanks to Ryan Wasserman for discussion here.) We set these sorts of worries aside, however, since our aim is to draw lessons for theorizing about moral responsibility.

about younger Zoe's moral responsibility. To secure this result, Cohen appeals to the following principle:

> *Intrinsic Mechanism*: whether a mechanism is moderately reasons-responsive depends only on the *intrinsic* properties of the agent in question.[9]

Cohen admits that Fischer & Ravizza do not explicitly endorse this principle, but he argues that it would be better, *ceteris paribus*, for them to accept it. And it certainly does have the ring of truth: after all, facts about the capacities of my decision-making processes don't seem to depend on anything happening across town. To know whether my capacities are reasons-responsive, it seems like you would only need to look at those capacities themselves.[10]

If we accept *Intrinsic Mechanism*, and we agree that the story of Zoe is metaphysically possible, then we can create a problem for Fischer & Ravizza. Recall that younger Zoe does not act from a moderately reasons-responsive mechanism, since there are no worlds in which she refrains from acting in self-defense, and hence no worlds that can serve as witness to the claim that her decision-making process is responsive to reasons. But now just tweak Zoe's story a bit, so that young Zoe does not face an older version of herself, but instead faces a time traveler that has no interesting counterfactual dependency on her – Cohen calls her 'Amy'. Notice that this tweak of the story doesn't alter any of *young Zoe's* intrinsic properties: all we have done is remove older Zoe from the story and replace her with a time traveler named Amy. But the second we break the counterfactual dependency between murderer and victim, we also get all the relevant possible worlds back in which young Zoe refrains from willing to nullify the attempted murder, which

[9] This is our paraphrase of Cohen's principle: "a moderately reasons-responsive mechanism *M* that issues in *S*'s φ-ing is *wholly* constituted by *S*'s intrinsic properties (either all of *S*'s intrinsic properties or, more likely, some subset thereof)" (2014: 2).

[10] While we can grant this claim for the sake of argument, Cohen's argument that Fischer & Ravizza should accept it is problematic. In particular, Cohen gives an example of one clearly irrelevant extrinsic property (being one mile away from a post office) and then claims that this suggests that only intrinsic properties are relevant to reasons-responsiveness. But this is a bit too quick; it wouldn't follow from the irrelevance of one extrinsic property that all extrinsic properties are irrelevant.

means that young Zoe miraculously becomes responsive to reasons again, despite our not having changed any of her intrinsic properties.

The upshot? If we accept *Intrinsic Mechanism*, then we have to say that young Zoe's mechanism is reasons-responsive in both stories or in neither, but the Fischer & Ravizza account is at odds with that verdict. According to Fischer & Ravizza's account, whereas young Zoe is not moderately reasons-responsive in the version of the story where she confronts her older self, young Zoe *is* moderately reasons-responsive in the version of the story where she confronts Amy (or, at least, there's no reason in the Amy story to think that young Zoe *isn't* moderately reasons-responsive). Something's gotta go, and since *Intrinsic Mechanism* is the most plausible of the bunch, the worry here can be adequately framed as a conflict between the metaphysical possibility of time travel and the Fischer & Ravizza account of moral responsibility.

3.

We've got three worries about Cohen's objection. The first worry shows that his objection, even if successful, is more limited in scope than it at first seems. The second two worries show that even the limited objection fails.

First: although Cohen describes his conclusion as the claim that Fischer & Ravizza's account of moral responsibility is incompatible with the metaphysical possibility of time travel, nothing quite so grand follows from the considerations he adduces, even if his arguments are sound. Rather, all that would follow is that *the time travel stories involving Zoe and Amy* are incompatible with the Fischer & Ravizza account of moral responsibility. Of course, we could generalize a further conclusion by abstracting away from the particular imaginary individuals in those stories, but still, at best that would give us the claim that the metaphysical possibility of

single-timeline backward time travel involving agents is incompatible with the Fischer & Ravizza account of moral responsibility. This is not an insignificant conclusion, since these are precisely the sorts of time travel stories that tend to capture the imaginations of sci-fi lovers. Still, single-timeline models of time travel are not the only feasible models, backward is not the only direction one might wish to travel, and, in the actual world at least, non-agential travel through time would probably be the first breakthrough to make headlines. So, Cohen's conclusion is more limited than advertised.

Even thus qualified, though, there are two major problems with Cohen's argument. The first is that Cohen doesn't respect the distinction that Fischer & Ravizza draw between *agents* and their *mechanisms*. The second is that Cohen fails to appreciate the significance of Fischer & Ravizza's claim that reactivity is "all of a piece", so that if a mechanism can react to *any* reason to do otherwise, then it can react to *all* such reasons.[11] We'll take these two problems in order.

First, consider one more time why younger Zoe seems not to be acting from a moderately reasons-responsive mechanism when she faces off against her older, time-traveling self. Although what Zoe *actually* does is will to nullify the attempted murder, in order for that to be an action for which she is morally responsible, there must be a suitable range of worlds in which Zoe recognizes reasons to refrain from nullifying the attempted murder, and there must be at least one world in which, having recognized those reasons, Zoe *does* refrain from nullifying the attempted murder. But since the would-be murderer is her older self, we know that there are *no* worlds in which she

[11] As an anonymous reviewer points out, if we consider a view like Fischer & Ravizza's but which lacks these two features (the distinction between agents and mechanisms and the claim that reactivity is all of a piece), such a view *would* fall prey to certain time travel scenarios (though Cohen's argument would still need to be qualified in the way we indicated above). But, as far as we know, no one holds such a view, and we are interested in defending Fischer & Ravizza's account. Perhaps, though, Cohen's challenge to Fischer & Ravizza serves to highlight the importance of these two features of the account.

refrains from nullifying the attempted murder. Hence, younger Zoe's nullifying actions can't have issued from a reasons-responsive mechanism.

But if you look closely at the justification just offered, you'll see that we've moved back and forth between talking about Zoe herself, on the one hand, and talking about Zoe's action-producing mechanism, on the other. And in fact, the justification gains whatever superficial plausibility it has precisely from this equivocation. On Fischer & Ravizza's official account, everything is done in terms of *mechanisms* rather than *agents*. So, in order to get the same result – that younger Zoe is not acting from a reasons-responsive mechanism when she nullifies her older self's attempted murder – we have to show, not that there are no worlds in which *Zoe* refrains from the act of nullifying, but rather that there are no worlds in which *her mechanism issues in an act of refraining*. It's the mechanism, after all, which has (or doesn't) the property of being responsive to reasons, and the agent acquires that status only derivatively.

Paying close attention to the difference between agents and mechanisms helps us to see how Fischer & Ravizza can escape Cohen's criticism. The feature of the time travel example that is so peculiar is that the person attempting murder and the person who is the victim of an attempted murder *are the same person* – this is why it doesn't make sense to imagine a world in which young Zoe fails to stop her own murder (i.e., a world in which she dies at the hands of her future self). But the Fischer & Ravizza account of moral responsibility doesn't apply at the level of persons – at least, not in the first instance. Instead, it applies at the level of *mechanisms*. And there's nothing contradictory about saying that the relevant mechanism might have issued in some other willing, since we need not hold everything fixed about the agent whose mechanism it is, in order to figure out what capacities the mechanism itself has.

Perhaps another way to put the point is to say that whereas there are no possible worlds in which young Zoe fails to stop *her own murder at the hands of her future self*, there certainly are possible worlds in which the type of mechanism on which young Zoe acts issues in the decision to let herself be killed – it's just that in those worlds, some of the external circumstances would have to be different. In those circumstances – the ones that we look to in order to figure out whether young Zoe's actually operative mechanism is responsive to reasons – perhaps the person attempting to murder her is an enemy combatant in a war, and she willingly sacrifices herself for the good of her community. There is, after all, no contradiction in the supposition that the mechanism on which young Zoe acts when she thwarts her older self's plan might nevertheless be the same kind of mechanism that, in a different circumstance, issues in a decision to sacrifice herself. (It's not as though young Zoe is *invincible*, after all.)

So, to sum up our first response to Cohen's objection: although there are no worlds in which young Zoe allows her older self to murder her, there are (it seems) plenty of worlds in which the relevant action-producing mechanism issues in a self-sacrificial decision due to the presence of different incentives. And it's this latter fact that tells us something about Zoe's moral responsibility, according to Fischer & Ravizza.

4.

The second reason why Cohen's objection fails has to do with a rather peculiar claim that Fischer & Ravizza make about the notion of *weak reasons-reactivity*. If you look back at the account of guidance control that we spelled out above, you'll notice that guidance control involves both *receptivity* and *reactivity*, but whereas Fischer & Ravizza classify the relevant sort of receptivity as *regular*, they classify the relevant sort of reactivity as *weak*. And indeed, when they spell out

what those terms mean, we can see that they correspond to different spheres of possible worlds. A mechanism is *regularly receptive* to reasons just in case there is an intelligible pattern of counterfactual circumstances in which the mechanism would "see" the reasons at play, but a mechanism is *weakly reactive* to reasons just in case there is *at least one* counterfactual circumstance in which the mechanism would respond to those reasons, upon seeing them. Why the asymmetry?

Fischer & Ravizza opt for *weak* reasons-reactivity because, as they put it, reactivity is "all of a piece" (1998: 73). Here's what they mean: "if an agent's mechanism reacts to *some* incentive to (say) do otherwise than he actually does, this shows that the mechanism *can* react to *any* incentive to do otherwise" (1998: 73). This is meant to mark a crucial difference between receptivity and reactivity. When it comes to receptivity, Fischer & Ravizza are worried about the possibility of a responsibility-undermining sort of blind spot in moral reasoning. They think it is possible, for example, that you might be able to recognize the fact that your action would break a promise as a reason not to do it, and yet you might not be able to recognize the fact that your action would cause me pain as a reason not to do it. That is, they are worried about mechanisms that are pathological in such a way that although certain moral reasons are on their radar, other moral reasons that seem like they should be equally visible just aren't on their radar. Such a person, Fischer & Ravizza maintain, ought to be excused due to this bizarre malfunction in receptivity.

But when it comes to "the capacity to *translate* reasons into choices (and then subsequent behavior)" (Fischer & Ravizza 1998: 69) – that is, when it comes to the capacity that Fischer & Ravizza call "reactivity" – their claim is that such a bizarre sort of "blind spot" is impossible. In fact, it wouldn't even be right to call it a "blind spot" in this instance, since we're talking about *reactivity* rather than *receptivity*. So, the "all of a piece" claim is that, so long as your mechanism

would react *at all* – so long as it's "online", so to speak – then it doesn't matter what precise reason we put into the mechanism. If there's a scenario in which it reacts to *one* reason, then it's got the capacity to react to them *all*. And that's why we only need to look at *one* possible world to determine whether a mechanism is appropriate *reactive* to reasons, even though we need to look at a suitably wide range of worlds to determine whether a mechanism is appropriately *receptive* to reasons.[12]

We've tried to keep the details to a minimum here, but they are important for seeing where Cohen's criticism goes wrong. Recall again why we are supposed to think that young Zoe fails to meet the criteria for exercising guidance control: given the peculiarities involved in backward time travel, there is no world in which young Zoe fails to stop her own murder, and this shows us that the mechanism on which she acts is insensitive to reasons. This is a point about *reactivity*: there is no possible scenario in which the relevant mechanism *reacts* to the reasons there may be for refraining from nullifying the murderous action, since "reaction" is a matter of translating reasons into choices and behavior, and of course it's not possible for young Zoe to be killed by her older self. So, it looks as though Zoe's mechanism doesn't have the sort of reactivity that Fischer & Ravizza think is needed for guidance control.

But again, this reasoning relies on a sort of equivocation. This time the equivocation is not between *agent* and *mechanism*, but instead between two sorts of *reason* to which the mechanism might react. If we focus just on the *exact reason* that young Zoe acts on – namely, one that makes essential reference to the peculiar situation she finds herself in, where her older self is trying to kill

[12] As an anonymous reviewer points out, Fischer & Ravizza's claim that reactivity is all of a piece seems to count as morally responsible some extremely weak-willed agents (e.g., a severe drug addict) who are intuitively not morally responsible, since there may well be one, possibly outlandish, scenario in which even a weak-willed agent's mechanism reacts successfully. Fischer & Ravizza explicitly acknowledge this implication of their view in their discussion of Brown and the drug "Plezu" (1998: 73-74). In later work, responding to an objection from Mele (2000), Fischer (2012: 187-192) tentatively suggests that we might say that such an agent is morally responsible but not blameworthy. For further discussion, see Cyr (2017: 315).

her – then Cohen is right to say that there is no possible scenario in which the mechanism reacts differently to *that reason*. However, this is not enough to show that the mechanism fails to satisfy the reactivity criterion on guidance control, because remember: according to Fischer & Ravizza, reactivity is *all of a piece*. So long as there is *at least one* possible scenario in which young Zoe's mechanism successfully reacts to a reason *of the same sort* as the one we are wondering about, then that's sufficient for us to conclude that the mechanism is *capable* of reacting to the reason we are wondering about.

In order to figure out whether young Zoe's mechanism is appropriately reactive to reasons, then, we don't need to find a scenario in which she fails to stop *her older self* from killing her. Instead, we just need to find a scenario in which she fails to stop *someone* from killing her. We just need to know whether the reason in question is the *sort* of reason her mechanism is able to translate into action; not whether there's a genuine possibility that *this particular* reason get translated into action.

Cohen considers an objection along these lines, that perhaps all we need to know about the mechanism is that it is capable of reacting to a threat from some "different but qualitatively similar" person. Cohen's response is to say that "even if there is a nomologically identical world in which Young Zoe refrains…from nullifying the act of someone who is qualitatively similar to Old Zoe, this has no bearing upon whether [Young Zoe's mechanism] is moderately reasons-responsive" (2014: 5). But this response fails to appreciate the claim that reactivity is "all of a piece". This claim is precisely what allows us to move from "possibly, young Zoe's mechanism reacts to a reason of the same sort" to "actually, young Zoe's mechanism is capable of reacting to the actual reason".[13]

[13] For discussion of Fischer & Ravizza's claim that reactivity is "all of a piece", and for a potential worry given that receptivity is *not* "all of a piece", see Todd & Tognazzini (2008).

5.

The last two sections have gotten us pretty far into the weeds, and that's because Cohen's objection focuses specifically on the Fischer & Ravizza theory of moral responsibility, which has been worked out in great detail. To construct an adequate response on behalf of Fischer & Ravizza, then, we've had to look at those details. But now we want to zoom out a bit. First, we'll give a high-altitude summary of why Cohen's objection fails. But then we'll try to articulate what we think is insightful about Cohen's worry, and what implications that insight has for theorizing about moral responsibility more generally. In the end we'll see that this will help us to bring even Fischer & Ravizza's account into sharper focus.

So, first, the high-altitude summary of our reply to Cohen. Cohen's basic worry is as follows: if moral responsibility is a matter of reasons-responsiveness, then merely changing an agent's external circumstances shouldn't make a difference to whether they are morally responsible. But cleverly constructed time travel examples can screw up counterfactuals about an agent without changing anything intrinsic to the agent herself, so if we understand reasons-responsiveness in terms of counterfactuals, then we'll be able to eliminate moral responsibility merely by changing an agent's external circumstances. Hence there is a deep tension at the heart of the Fischer & Ravizza account. On the one hand, they want reasons-responsiveness to be a matter of an agent's *intrinsic* properties, but on the other hand, they want to understand reasons-responsiveness in terms of *counterfactuals*. And what time travel stories show us (among other things) is that counterfactuals about an agent can vary independently of the agent's intrinsic properties, so it looks like Fischer & Ravizza can't have both of the things they want.

Our basic reply is to say that Cohen has been looking at the wrong counterfactuals. Time travel examples involving retro-suicide attempts do mess up counterfactuals about the *agent*, but it's not clear that they mess up counterfactuals about the *mechanism*. (This was our first substantive reply.) Moreover, even if time travel examples show that there's no way the mechanism will react to the *actual reason*, that doesn't show that the mechanism *can't* react to the actual reason, since reactivity is all of a piece. All Fischer & Ravizza need is the claim that there is some reason of the same sort that the mechanism possibly reacts to. (This was our second substantive reply.)

But even if Cohen's objection fails, there is likely to be a lingering worry here, which might be expressed rhetorically as a question: why exactly is an *actual-sequence* account of moral responsibility trafficking in *counterfactuals* in the first place? Facts about what *could* have or *would* have happened seem like the basic ingredients of a theory of moral responsibility that emphasizes *alternative possibilities*. True, Fischer & Ravizza make the move from talking about what an *agent* can do to talking about what a *mechanism* can do (or is capable of doing), but this move might seem a bit like cheating, since it seems to smuggle alternative possibilities in through the back door.[14] Cohen's objection is made possible by the fact that Fischer & Ravizza emphasize the importance of counterfactuals, and yet we can use time travel to generate some surprising counterfactual results. Although the objection fails, it provides the occasion to rethink the framing of Fischer & Ravizza's view, since – in our view – counterfactuals ought not to have a prominent place in an actual-sequence theory of moral responsibility in the first place.

[14] It has seemed that way to many commentators. See, for example, Watson (2001: 382).

6.

We're not the first to note the awkwardness of being committed to an *actual-sequence* account of moral responsibility but yet giving pride of place to *counterfactuals* in the details of that theory. This criticism has also been raised forcefully by Christopher Franklin (2015), in his descriptively titled paper, "Everyone Thinks that an Ability to Do Otherwise is Necessary for Free Will and Moral Responsibility". According to Franklin, despite their claim to be providing an actual-sequence account of moral responsibility, Fischer & Ravizza's account requires alternative possibilities after all. As we have seen, Fischer & Ravizza's account of guidance control includes the following reactivity component:

> A mechanism is *weakly reactive to reasons* just in case it is regularly receptive to reasons and, in at least one of the possible scenarios described in the account of regular receptivity, the agent chooses and does otherwise for the reason in question.

In order for an agent to be morally responsible, then, the agent's operative mechanism must react to a reason to do otherwise in some possible scenario. But this is just to say that the mechanism can do (or is capable of doing) otherwise, which is tantamount to saying that the mechanism has alternative possibilities. In Franklin's words, Fischer & Ravizza are committed to the view that "a mechanism is appropriately reactive only if it has certain dispositions or abilities, namely the ability to act on different sufficient reasons" (2015: 2096).

Again, in order to preserve the insight of Frankfurt-style counterexamples, Fischer & Ravizza aim to show that morally responsible agents need not be able to do otherwise, or have alternative possibilities, even though the account does require that morally responsible agents act from a weakly reactivity mechanism. As Franklin argues, however, what's true of agents' mechanisms holds for agents themselves too:

> Agents make choices, act, and are morally responsible in virtue of the activity of their mechanisms…If the agent's mechanism is able to do otherwise, then the agent is, in virtue

> of taking responsibility for the mechanism, able to do otherwise. A central contention, therefore, of Fischer [& Ravizza]'s theory of moral responsibility is that agents are morally responsible only if they possess an ability to do otherwise. (2015: 2097)

If Franklin is right, then why do Fischer & Ravizza deny that morally responsible agents must have the ability to do otherwise? Franklin says that it is because Fischer & Ravizza really intended (or at least should have intended) to say that "certain *species* of abilities are irrelevant" (2015: 2097, emphasis added), specifically the sort of ability that agents in Frankfurt-style counterexamples *lack*.[15] But once we distinguish that sort of ability from the ability required by the reactivity component of Fischer & Ravizza's account, it is clear that the account does require *some* ability to do otherwise.

Now, we think that Franklin's criticism fails because he has conflated an ability to do otherwise with the mere presence of "alternative possibilities".[16] It's true that Fischer & Ravizza look to possible worlds in order to determine whether an agent's mechanism is suitably reasons-responsive, but it does not follow from the modal facts themselves that an agent who acts from a suitably reasons-responsive mechanism is thereby *able* to have done otherwise. To have an ability requires more than the possession of just any alternative possibility. For example, it may be that getting a hole-in-one is a genuinely possible alternative to my hitting the bunker, but (trust me) I do not have the ability to hit a hole-in-one. Still, one might think that the spirit behind Franklin's criticism survives this response. The lesson we are supposed to have learned from Frankfurt-style counterexamples, one might think, is that facts about other worlds are simply *irrelevant* to whether an agent is actually morally responsible for what they do. And so there appears to be a sense in

[15] This is related to the distinction some authors draw between "general" and "specific" abilities. See, for example, Mele (2003) and Whittle (2010).
[16] A detailed version of this response to Franklin can be found in Cyr (2017). See also Kittle (2019).

which Fischer & Ravizza – those great champions of Frankfurt-style compatibilism – have misunderstood the central lesson of the examples.

But a lot depends here on what is meant by the term 'irrelevant'. As the literature on Frankfurt-style counterexamples and semicompatibilism developed late last century, the main question was whether an ability to do otherwise was *necessary* for moral responsibility. Actual-sequence theorists said "no", whereas leeway theorists said "yes". Over the last 20 years, however, philosophers have more carefully distinguished between "mere" necessary conditions of a claim, on the one hand, and factors *in virtue of which* a claim is true.[17] And that means that there are now *three* different views theorists might have on the question of how alternative possibilities relate to moral responsibility.

> *Necessary and Grounded In*: Someone's being morally responsible not only *entails* the presence of alternative possibilities, but is *partly grounded in* the existence of those alternative possibilities.
>
> *Necessary but Not Grounded In*: Someone's being morally responsible *entails* the presence of alternative possibilities, but it is *not even partly in virtue of* those alternative possibilities that the person is morally responsible.
>
> *Neither Necessary Nor Grounded In*: Someone's being morally responsible *neither entails nor is grounded* in facts about alternative possibilities.

Although Fischer & Ravizza were writing before the contemporary literature on grounding really took off, it is clear that their theory falls into the second of these three categories, and *this* is the sense in which it is an "actual-sequence" theory: although facts about other worlds follow from their account of reasons-responsiveness, it is not *in virtue of* those other-worldly facts that a mechanism is reasons-responsive.[18] Rather, those other-worldly facts are what they are *because*

[17] See, for example, Fine (1995), Correia (2008) and Clark & Liggins (2012).

[18] As a matter of historical interest, Frankfurt himself has clearly distinguished between necessary conditions for moral responsibility, on the one hand, and facts in virtue of which someone is morally responsible, on the other, and he agrees with Fischer here. Responding to a criticism of his argument against the Principle of Alternative Possibilities, Frankfurt (2003: 340) says, "The critical issue concerning PAP, then, is not whether it is always possible that an agent

the actual-world mechanism is reasons-responsive. It's easy to conflate "direction of reasoning" with "direction of explanation", but they are crucially different. The other-worldly facts are reasons to believe that the actual mechanism is reasons-responsive, but they are not explanations of why it has that feature.[19]

After making their claim that "reactivity is all of a piece" (discussed above), for example, Fischer & Ravizza appeal to grounding:

> Our contention, then, is that a mechanism's reacting differently to a sufficient reason to do otherwise in some other possible world shows that the same kind of mechanism can react differently to the *actual* reason to do otherwise. This general capacity of the agent's actual-sequence mechanism—and *not* the agent's power to do otherwise—is what helps to ground moral responsibility. (1998: 73)

In more recent work, Fischer has again made this point quite explicit, conceding to Franklin that perhaps he could have been clearer in previous work. Fischer says:

> I completely agree with Franklin that I do indeed believe that various kinds of alternative possibilities are required for moral responsibility (although not for the 'grounding' or explanation of moral responsibility), and thus that my repeated contention that alternative possibilities are not required for moral responsibility might well have caused confusion… But, as Franklin also notes, these were not the sorts of alternative possibilities I had in mind in contending that moral responsibility does not require alternative possibilities. I have absolutely no interest in showing that moral responsibility does not require general capacities or abilities to do otherwise, or various other kinds of abilities to do otherwise that abstract away from the particulars of the agent's history and/or present situation…I have always been interested in the sort of alternative possibility that would be (or could plausibly be thought to be) ruled out by causal determinism. And, clearly, general abilities and indeed any sort of ability to do otherwise that abstracts away from features of the

---

who is morally responsible for performing a certain action might have acted differently. Rather, it is whether that possibility – even assuming that it is real – *counts for anything* in determining whether he is morally responsible for what he did. My claim is that it does not" (our emphasis). See also Leon and Tognazzini (2010) for an examination of how the difference between necessity and grounding ought to shape our understanding of Frankfurt-style counterexamples.

[19] An anonymous reviewer points out that even if Fischer & Ravizza don't give the other-worldly facts a role in grounding an agent's responsibility, merely acknowledging that they follow from the presence of responsibility is enough to undermine Fischer & Ravizza's claim to be offering a semi-compatibilist account of moral responsibility. Semi-compatibilism is usually understood as the view that moral responsibility is compatible with determinism, regardless of whether determinism rules out the ability to do otherwise. But now if Fischer & Ravizza acknowledge that reasons-responsive mechanisms generate alternative possibilities, it looks like it *does* matter after all whether determinism rules out all alternative possibilities. But as we point out in the text just below, Fischer distinguishes the sort of alternative possibilities entailed by the presence of a reasons-responsive mechanism from the sort of ability to do otherwise that features in the official formulation of the semicompatibilist view.

> agent's past and/or current situation need not be inconsistent with causal determinism. (2018: 221)

So, even if Fischer's view implies that alternative possibilities are necessary for moral responsibility, and even if the view implies that *some* (general) ability to do otherwise is necessary for moral responsibility, Fischer maintains that these possibilities/abilities do not ground or explain moral responsibility.[20]

In this way, the theory of Fischer & Ravizza (as well as Fischer's more recent work) contrasts with two other sort of compatibilist views, the first of which takes the "neither/nor" option and the second of which takes the "both/and" option. Mesh theories like those inspired by Frankfurt (1971) and Watson (1975) offer accounts of moral responsibility according to which one need not even mention what happens in other worlds. Frankfurt himself is explicit, in fact, that moral responsibility does not require reasons-responsiveness:

> I do not believe that the mechanism has to be reasons-responsive. The mechanism is constituted by desires and volitions and, in my view, what counts is just whether what the agent wills is what he really wants to will…Someone who is wholeheartedly behind the desires that move him when he acts is morally responsible for what he does, in my judgment, whether or not he has any reasons for his deeds or for his desires. (Frankfurt 2002: 28).

Now, perhaps a comprehensive account of "wholeheartedness" would need to appeal to other-worldly facts; we don't intend to take a stand on how best to spell out a mesh theory of the sort inspired by Frankfurt's work. The point is simply that, at least on the face of it, a mesh theory looks to be even more of an "actual-sequence" theory than a theory that emphasizes reasons-responsiveness. Whereas reasons-responsiveness theories entail facts about what agents are up to

[20] Carolina Sartorio (2016) also opts for a version of compatibilism according to which facts about possible worlds are necessary but not part of what grounds an agent's moral responsibility. Sartorio goes one step further than Fischer & Ravizza, though, and claims that the other-worldly facts show us that *absences* are playing a causal role in the actual sequence.

in other worlds, it's not clear that mesh theories do. They are similar, however, in rejecting the idea that an agent's moral responsibility is even partly *grounded* in those other-worldly facts.

However, there are compatibilist theories that take a "both/and" approach instead. Here we have in mind the view of the so-called "new dispositionalists", who not only reject Frankfurt-style counterexamples, but who also aim to give a positive view of free will in terms of dispositions, which are spelled out in counterfactual terms.[21] These are leeway compatibilists rather than source compatibilists, theorists who think that not only is an ability to do otherwise necessary for moral responsibility, but also that one is morally responsible partly in virtue of such an ability. Even if Franklin is right that reasons-responsive theorists are aligned in an important way with leeway theorists – since they both develop theories that give pride of place to facts about other worlds – there is nevertheless an important difference between them – since one seeks to explain moral responsibility in terms of those other-worldly facts, whereas the other seeks to explain moral responsibility only in terms of actual-sequence facts.

Fittingly, then, we have found another way in which the theory of Fischer & Ravizza is a *semicompatibilist* theory. The familiar sense of that term conveys the idea that determinism is compatible with moral responsibility, regardless of whether determinism rules out the ability to do otherwise. But now we've seen that Fischer & Ravizza also hold the view that moral responsibility is not even partly grounded in the presence of alternative possibilities, regardless of whether the facts that ground moral responsibility entail the existence of alternative possibilities. The first claim differentiates Fischer & Ravizza from leeway compatibilists like Vihvelin, whereas the

[21] See, for example, Vihvelin (2004, 2013) and Fara (2008). For a critique of these accounts, see Clarke (2009) and Franklin (2011).

second claim differentiates them from what we might say are "pure" actual-sequence compatibilists, such as Frankfurt.[22]

7.

We have seen that attending to the distinction between necessity and grounding has not only clarified Fischer & Ravizza's view but has also provided a clearer view of how it differs from rival actual-sequence approaches as well as from alternative-possibilities approaches. In conclusion, let's briefly return to Cohen's argument against Fischer & Ravizza from the possibility of time travel. We are now in a better position to appreciate why it seemed appealing in the first place, despite its unsoundness.

Recall Cohen's story: younger Zoe responds to older Zoe in self-defense, and there is no world in which Zoe refrains from acting in self-defense since older Zoe's existence depends counterfactually on young Zoe's responding in self-defense. Cohen takes this case to raise a problem for Fischer & Ravizza since younger Zoe seems not to be responsive to reasons, on their account, and yet an intrinsic duplicate of younger Zoe could be responsive to reasons in different circumstances (where the self-defense is in response to someone whose existence does not depend counterfactually on Zoe's response). Crucially, the problem is that there do not seem to be any differences in the grounds of young Zoe's moral responsibility from one case to the next, despite the difference in facts about their alternative possibilities. In other words, the case of time travel

[22] A wrinkle worth noting but not worth dwelling on: there is room for a theory of moral responsibility according to which (1) the ability to do otherwise is part of the explanation for why someone is morally responsible, and (2) the ability to do otherwise is not to be analyzed in terms of counterfactuals, but instead is to be taken as more fundamental than the counterfactuals it supports. This sort of theory would resemble Fischer & Ravizza's in that moral responsibility is fully explained by facts about the actual-sequence, yet it would differ from Fischer & Ravizza's in appealing to an ability to do otherwise. Fischer & Ravizza are interested in distancing themselves from those two sorts of theorists: those who think the ability to do otherwise is required for moral responsibility, and also those who think that facts about other worlds are part of what grounds moral responsibility. What we are pointing out here is that those two sets of theorists are disjoint.

that features in Cohen's objection to Fischer & Ravizza allows us to falsify counterfactuals about young Zoe without altering any of the actual-sequence facts about young Zoe's moral competence that ground her moral responsibility.

We have argued that Cohen's argument is unsound, but there's an important lesson to learn from the argument nevertheless, which is that actual-sequence compatibilists ought to de-emphasize, or at least properly contextualize, the role that counterfactuals play in their theories. To the extent that it seems like those counterfactuals are doing the work of *grounding* an agent's moral responsibility, the theory will seem vulnerable to the sort of objection that Cohen launches. Whatever reasons-responsiveness is, it needs to be conceived as something that generates its associated counterfactuals rather than being constituted or constrained by them.[23]

References

Arntzenius, Frank & Maudlin, Tim. (2002) "Time Travel and Modern Physics," in C. Callender, ed., *Time, Reality and Experience* (Cambridge: Cambridge University Press: 169-200).

Clark, Michael & Liggins, David. (2012) "Recent Work on Grounding," *Analysis* 72: 812–823.

Clarke, Randolph. (2009) "Dispositions, Abilities to Act, and Free Will: The New Dispositionalism," *Mind* 118: 323-351.

Cohen, Yishai. (2014) "Reasons-Responsiveness and Time Travel," *Journal of Ethics and Social Philosophy* 8: 1-7.

Correia, Fabrice. (2008) "Ontological Dependence," *Philosophy Compass* 3: 1013-1032.

Cyr, Taylor W. (2017) "Semicompatibilism: No Ability to Do Otherwise Required," *Philosophical Explorations* 20: 308-321.

Dowe, Phil. (2000) "The Case for Time Travel," *Philosophy* 3: 441-51.

[23] Thanks very much to Ryan Wasserman and two anonymous reviewers for helpful comments on a previous version of this paper.

Fara, Michael. (2008) “Masked Abilities and Compatibilism”, *Mind* 117: 844-865.

Fine, Kit. (1995) “Ontological Dependence,” *Proceedings of the Aristotelian Society* 95: 269–290.

Fischer, John Martin & Ravizza, Mark. (1998) *Responsibility and Control: A Theory of Moral Responsibility*. New York: Cambridge University Press.

Fischer, John Martin. (2006) *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.

Fischer, John Martin. (2012) *Deep Control: Essays on Free Will and Value*. New York: Oxford University Press.

Fischer, John Martin. (2018) “The Freedom Required for Moral Responsibility,” in D. Brink, S. S. Meyer, & C. Shields (eds.), *Virtue, Happiness, Knowledge: Themes from the Work of Gail Fine and Terence Irwin* (New York: Oxford University Press): 216-233.

Frankfurt, Harry. (1969) “Alternate Possibilities and Moral Responsibility,” *Journal of Philosophy* 66: 829-839.

Frankfurt, Harry. (1971) “Freedom of the Will and the Concept of a Person,” *Journal of Philosophy* 68: 5-20.

Frankfurt, Harry. (2002) “Reply to John Martin Fischer,” in S. Buss and L. Overton (eds.), *Contours of Agency: Essays on Themes from Harry Frankfurt* (Cambridge, MA: MIT Press): 27-32.

Frankfurt, Harry. (2003) “Some Thoughts Concerning PAP,” in D. Widerker and M. McKenna (eds.), *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (Aldershot: Ashgate): 339-346.

Franklin, Christopher Evan. (2011) “Masks, Abilities, and Opportunities: Why the New Dispositionalism Cannot Succeed”, *The Modern Schoolman* 88: 89-103.

Franklin, Christopher Evan. (2015) “Everyone Thinks that an Ability to Do Otherwise Is Necessary for Free Will and Moral Responsibility,” *Philosophical Studies* 172: 2091-2107.

Kittle, Simon. (2019) “Does Everyone Think the Ability to do Otherwise is Necessary for Free Will and Moral Responsibility?” *Philosophia* 47: 1177-1183.

Leon, Felipe and Neal A. Tognazzini. (2010) “Why Frankfurt-Examples Don’t Need to Succeed to Succeed,” *Philosophy and Phenomenological Research* 80: 551-565.

Lewis, David. (1976) “The Paradoxes of Time Travel,” *American Philosophical Quarterly* 13: 145–152.

McCormick, Kelly. (2017) “A Dilemma for Morally Responsible Time Travelers,” *Philosophical Studies* 174: 379-389.

McKenna, Michael. (2013) “Reasons-Responsiveness, Agents, and Mechanisms,” in D. Shoemaker (ed.), *Oxford Studies in Agency and Responsibility*, vol. 1 (Oxford: Oxford University Press): 151-183.

Mele, Alfred R. (2000) “Reactive Attitudes, Reactivity, and Omissions,” *Philosophy and Phenomenological Research* 61: 447–452.

Mele, Alfred R. (2003) “Agents’ Abilities,” *Nous* 37: 447–470.

Sartorio, Carolina. (2016) *Causation and Free Will*. New York: Oxford University Press.

Spencer, Joshua. (2013) “What Time Travelers Cannot Not Do (but Are Responsible for Anyway),” *Philosophical Studies* 166: 149–162.

Todd, Patrick & Tognazzini, Neal A. (2008) “A Problem for Guidance Control,” *Philosophical Quarterly* 58: 685-692.

Vihvelin, Kadri. (1996) “What Time Travelers Cannot Do,” *Philosophical Studies* 81: 315-330.

Vihvelin, Kadri. (2004) “Free Will Demystified: A Dispositional Account,” *Philosophical Topics* 32: 427-450.

Vihvelin, Kadri. (2013) *Causes, Laws, and Free Will: Why Determinism Doesn’t Matter*. New York: Oxford University Press.

Wasserman, Ryan. (2018) *Paradoxes of Time Travel*. New York: Oxford University Press.

Watson, Gary. (1975) “Free Agency,” *Journal of Philosophy* 72: 205-220.

Watson, Gary. (2001) “Reasons and Responsibility,” *Ethics* 111: 374-394.

Whittle, Ann. (2010) “Dispositional Abilities,” *Philosophers’ Imprint* 10: 1–23.