CrossMark

Freedom, self-prediction, and the possibility of time travel

Alison Fernandes[1] (iD)

Abstract Do time travellers retain their normal freedom and abilities when they travel back in time? Lewis, Horwich and Sider argue that they do. Time-travelling Tim can kill his young grandfather, his younger self, or whomever else he pleases— and so, it seems can reasonably deliberate about whether to do these things. He might not succeed. But he is still just as free as a non-time traveller. I'll disagree. The freedom of time travellers is limited by a *rational constraint*. Tim can't *reasonably* deliberate on killing his grandfather, certain that he'll fail. If Tim follows his evidence, and appropriately self-predicts, he will be certain he won't kill his grandfather. So if Tim is both evidentially and deliberatively rational, he can't deliberate on killing his grandfather. This result has consequences. Firstly, it shows how evidential limits in the actual world contribute to our conception of the future as open. Secondly, it undercuts arguments against the possibility of time travel. Thirdly, it affects how we evaluate counterfactuals in time travel worlds, as well as our own. I'll use the constraint to motivate an evidential and temporally neutral method of evaluating counterfactuals that holds fixed what a relevant deliberating agent has evidence of, independently of her decision. Using this method, an agent's local abilities may be affected by what happens globally at other times, including the future.

Keywords Time travel · Freedom · Deliberation · Causation · Counterfactuals · Time asymmetry

---

✉ Alison Fernandes
alison.fernandes@tcd.ie

1   Department of Philosophy, Trinity College Dublin, College Green, Dublin 2, Ireland

1 Introduction

Tim travels back in time, buys a shotgun and takes aim at his young grandfather. He intends to kill young-gramps, before young-gramps has a chance to grow up, meet Tim's grandmother, and make the family fortune in munitions. Can Tim kill young-gramps? David Lewis (1976), Paul Horwich (1987) and Ted Sider (2002) argue that Tim *can* kill his young grandfather, his younger self, or whomever else he pleases. True, Tim won't actually succeed in killing young-gramps. Leaving aside regeneration, branching time, and the like, Tim's presence in that past implies that he fails—gramps survives. But, they claim, what prevents Tim succeeding is not that he lacks the ability to kill gramps. Some everyday sort of failure will occur—the gun jams, Tim's nerve fails, he slips on a banana-peel. 'His failure by no means proves that he was not really able to kill Grandfather' (Lewis 1976, p. 150).[1] Nor do other features of the case affect Tim's freedom. Tim's knowledge that he will fail 'in no way detracts from his ability to have behaved otherwise than he did' (Horwich 1987, p. 117). Even if Tim's attempts are necessarily accompanied by an 'apparently miraculous series of coincidences', 'this fact is unremarkable, and in no way undermines time travelers' freedom' (Sider 2002, p. 135). Agents retain their normal freedom and abilities when they travel back in time.

What prevents Tim succeeding may well be an everyday sort of failure. But, I'll argue, there is nevertheless an important way in which Tim's freedom *is* compromised. Why might Tim be less than free? Vihvelin (1996) argues in an analogous case that Tim is not free to kill young-gramps, because it is false that, were Tim to try to kill young-gramps, he might succeed. I'll put such counterfactuals to the side for now (and return to them in Sect. 6). There are other reasons for thinking Tim's freedom is compromised. According to an ignorance condition on deliberation, agents can't reasonably deliberate when they're already certain of what they will or won't do—when they 'self-predict' their own behaviour (or its intended results). According to a distinct evidential norm, agents should be certain of what their evidence settles. If Tim time travels to the past and retains his memories, he has evidence that settles he won't succeed at killing young-gramps. If Tim's beliefs conform to his evidence, he'll be certain he won't succeed. If so, Tim can't reasonably deliberate on killing young-gramps. Insofar as time travellers like Tim are evidentially and deliberatively rational, they won't deliberate on parts of the future we would expect them to control in the actual world.

Why does this result matter? Firstly, it shows how evidential limits in the actual world contribute to our conception of the future as open. An aspect of the future's apparent openness is that it appears open to reasonable deliberation. According to the constraint, part of the reason why we can reasonably deliberate on the future, while following our evidence, is the fact that we *don't* have records of the future. Secondly, as I'll explain, the result undercuts arguments against the possibility of

---

[1] While Lewis (1976, p. 150) takes the evaluation of abilities to depend on context, he implies that the only contexts in which Tim's failure are held fixed are fatalist ones, and don't concern what Tim can do in any reasonable sense.

time travel, and shows why cases of time travel will produce fewer apparently miraculous coincidences. Thirdly, the constraint matters to how we evaluate counterfactuals and abilities in the actual world. I'll use the constraint to motivate an evidential and temporally neutral method of evaluating counterfactuals—one that holds fixed what the relevant deliberating agent has evidence of, independently of her decision.

A consequence of these arguments is that rational constraints on deliberation matter to what may seem like distinctly *metaphysical* issues: the apparent openness of the future, the possibility of time travel, and the evaluation of counterfactuals and abilities. While perhaps surprising, this result fits well with recent work in areas like the philosophy of causation. Agent-based (Price 1991; Ismael 2007; Blanchard 2014; Fernandes 2017) and interventionist accounts (Pearl 2000; Woodward 2003) take the usefulness of causal information to agents to be central to understanding causation. While causal relations appear when agents aren't around, causal relations are significant because of our interest in shaping the world. When evaluating counterfactuals and abilities in situations involving complex causal structures, like time travel cases, there is thus good precedent for attending to the interests of deliberating agents.

Let me note some assumptions. I assume that backwards time travel is metaphysically possible, and that Lewis' response (1976) to the grandfather paradox is correct—time travel does not lead to contradictions. I assume that time travel involves backwards causation (Lewis 1976), rather than branching time or multiple time lines. I'll *not* assume that time travel operates by any particular mechanism, such as closed timelike curves. While such an assumption is useful for exploring the solution space of General Relativity, it's not needed to explore the relations between causation, counterfactuals and abilities.[2] I'll assume a broadly 4-dimensionalist view of time, such that talk of the present, past and future is to be treated indexically. The referent of 'the present' will shift as required by context.

The paper proceeds as follows. In Sect. 2, I present an ignorance condition on deliberation, and show how it constrains a time traveller's deliberative freedom. In Sect. 3, I defend the condition from objections. In Sects. 4, 5 and 6 I explore how the constraint matters to broader metaphysical issues, including explaining the apparent openness of the future (Sect. 4), undercutting arguments against the possibility of time travel (Sect. 5), and developing an evidential and temporally neutral method of evaluating counterfactuals (Sect. 6). While some readers may be most interested in these results, the constraint is controversial enough that I will spend some time in its initial presentation and defence (Sects. 2, 3).

---

[2] Even if spacetimes involving closed timelike curves allow for no global time ordering, so that we can only define temporal direction locally, and backwards time travel doesn't involve *local* backwards causation, we still face problems evaluating counterfactuals that aren't local—see Sect. 6.

2 How self-prediction constrains deliberation

Tina is deliberating about whether to take up strength training. She considers its costs and benefits, and decides that she will. Take practical deliberation of this form to be a decision-making process, in which an agent deliberates between different options she takes to be available, and which aims to issue in decisions about what action to perform, or what option is to result. A plausible ignorance condition on such deliberation is that Tina can't reasonably deliberate about whether to take up strength training, if she is already certain that she will take up strength training, or certain that she won't. If she 'self-predicts' her behaviour in this way, her deliberation is, in an important sense, unnecessary and out of place—she may as well turn her attention to other things. In more general terms, an agent can't reasonably deliberate about whether to ø, if she's certain she will ø, or certain she won't ø (where ø is an action).[3] Practical deliberation requires uncertainty.

My aim is not to defend this ignorance condition in full. It is enough for my purposes that it's sufficiently plausible to make its consequences worth exploring. Why is such a condition plausible? Firstly, it helps explain why certain possibilities are ruled out as live options. It seems Tina shouldn't deliberate on the sun rising tomorrow, or on $2 + 2$ equaling 4, or on giving up smoking if she is certainly she'll fail. If Tina can only reasonably deliberate on epistemic uncertainties, we can explain why these possibilities aren't deliberative alternatives for her. Secondly, the ignorance condition is motivated by considering the doxastic and epistemic functions of deliberation. Regarding its *doxastic* function, deliberation is a way of settling how we'll act so that we can engage in further planning (Bratman 1984). Once Tina has decided she will take up strength training, she can plan her program, or find a trainer. Having decided on *A*, 'in any additional theoretical or practical reasoning, one will be able to take it for granted that one will be doing A' (Harman 1976, p. 438). For similar arguments, see Grice (1971), Levi (1986, p. 86), Holton (2006) and Fernandes (2016b, chap. 2). If deliberation is to serve this doxastic function, and allow agents to come to certainty by making a decision, agents must be uncertain as they deliberate. Regarding its *epistemic* function, deliberating can be a way of coming to *know* what one will do, by its output—decision. If agents have other appropriately justified beliefs, decision-making is a way of gaining knowledge. In Ginet's terms, 'the whole point of making up one's mind is to pass from uncertainty to a kind of knowledge about what one will do or try to do' (1962, p. 52). Unless agents have unusual belief sets, coming to know what one will do requires being uncertain first. Even if practical deliberation has other important benefits, such as helping us make *good* decisions, and even if reasons play an *essential* role in deliberation, the centrality of deliberation's doxastic and epistemic functions provides good reason to think that an ignorance condition must be satisfied for reasonable deliberation. Ignorance conditions against self-prediction

---

[3] I won't distinguish sharply between actions and their intended results. For some purposes, a more general formulation is preferable, one that uses de-tensed propositions or states of affairs, so as not to presume agents deliberate only on the future. But, for grammatical ease, I'll use the future-tensed action formulation.

have been defended as conditions on forming intentions (Harman 1976; Rennick 2015), forming decisions (Ginet 1962), recognising an option as available in deliberation (Levi 1986, Chap. 4; Kapitan 1986), being 'epistemically free' in deliberation (Fernandes 2016a), and on deliberating practically (Hampshire and Hart 1958; Dennett 1984, p. 113). For further discussion of different ways that epistemic conditions can feature in accounts of deliberation, see Fernandes (2016a).

The ignorance condition concerns when deliberation is *reasonable*, not when it is *possible simplicter*. Whether deliberation is possible depends on how closely one ties psychological possibility to rationality—that is, how normatively one characterises mental states. A middle-ground view is that agents may fail to satisfy the ignorance condition on particular occasions, but must satisfy it in general if they are to count as deliberating. I won't assume a particular view about psychological possibility in what follows, although I will later rely on the assumption that agents can be *expected* to be rational for the most part (Sect. 5). Self-predicting agents may still be able to engage in other forms of practical reasoning, consistent with the ignorance condition. Tina may be able to consider her reasons for taking up strength training, how she might do so, or whether she *ought* to. I won't take a stance on how deeply into practical reasoning the ignorance condition should extend. Finally, note that the condition is *necessary*, and *not sufficient*, for reasonable deliberation. For sufficiency, other conditions will be needed as well—see Fernandes (2016b, chap. 2) for some suggestions.

There are tricky cases. What if Tina *decides* to take up strength training, and thereby becomes certain that she will. Does this mean she can no longer reasonably deliberate on whether to? My own view is that Tina can only reasonably deliberate if she reneges on her decision or otherwise becomes uncertain (Owens 2008, p. 263). But these cases aren't the ones I'll be concerned with. I'll only consider cases where an agent's certainty *doesn't* arise by her making or inferring from her decision. The ignorance condition can be weakened to the following:

Ignorance   An agent can't reasonably deliberate on whether to ø if she is certain that she will ø or certain that she won't ø, and (in either case) this certainty does not arise from her making or inferring from her decision.

One might worry that this condition is too weak to be interesting. Perhaps agents can never be *certain* of what they'll do. Let me clarify the notion of certainty involved. Say you leave the conference room for lunch. You presume that the room will be there when you return. When pressed, you might admit that you're not 100% sure of this. After all, earthquakes happen from time to time. Nor are you willing to bet your life savings on it. You might not have a 'degree of belief' 1 in the room being there. Yet, for relevant practical and theoretical purposes, concerning where you leave your bag, and answering questions like "Where will the conference be?," you take the room's after-lunch existence to be settled—it is not something you trouble yourself concerning. The state of affairs is a fixed point in what you take to be the case, such that you can make further plans without having to reconsider it.

This is the notion of certainty I have in mind: 'practical certainty'.[4] An agent is practically certain when a given state is settled by her beliefs, relative to the practical and theoretical contexts relevant to her.

Let's return to time-travelling Tim. Can Tim reasonably deliberate on killing young-gramps? If Tim thinks clearly about the matter, he will be certain he won't succeed at killing young-gramps. Why? According to the standard story, Tim deliberately hunts down his young grandfather. Tim is certain he takes aim at young-gramps, rather than some hapless stranger. Tim is equally certain that young-gramps grows up to meet grandmother, make the family fortune in munitions, and live to an old age. Indeed, Tim's having these beliefs partly explains *why* he travels back in time.[5] If Tim puts these pieces together, he'll be certain that young-gramps survives to an old age, and that he won't succeed in killing young-gramps. If Tim is certain he won't kill young-gramps, then, by *Ignorance*, he can't reasonably deliberate on killing young-gramps.[6] His self-prediction rules out reasonable deliberation.

Stephanie Rennick (2015) argues for a similar conclusion. There are intentions that sufficiently reflective time travellers cannot form. Rennick presents this conclusion as a contrast between ordinary person-on-the-street time travellers and the more philosophically inclined: there are 'Things mere mortals can do, but philosophers can't' (ibid. p. 22). It's certainly true that Tim may be unreflective, and so uncertain about whether he'll succeed at killing young-gramps—perhaps because he's unsure about the correct metaphysics of time (see Objection 5). But there is, nevertheless, an important sense in which Tim *should* be certain. A part of the standard setup is that Tim has evidence that he won't succeed, in the following sense: Tim has access to local states of affairs now that are reliably correlated with young-gramps' survival. Tim has memories of his grandfather living to an old age, and evidence that the man he confronts is his grandfather. Indeed, Tim can travel back in time with a vast quantity of photographs, written testimony, family letters, newspaper clippings and other records that reliably indicate young-gramps' survival. The causal structure of time travel allows Tim to have records of states of affairs that lie in the future of his encounter with young-gramps—giving Tim evidence that settles that he won't succeed.[7]

Let me clarify this somewhat non-standard notion of 'evidence'. The notion derives from thinking about the public character of evidence in courts of law, medicine, science, and the everyday, where evidence is something we can *show* one another. A piece of evidence is a local state now that is reliably correlated with another state by causal or other nomic means, and that thereby counts as

---

[4] For related notions, see Levi (1980, 1986), Harman (1976, p. 438) and Dennett (1984, p. 113).

[5] Admittedly, it's difficult to make sense of Tim's motivation for travelling back without taking him to be irrational, or to have adopted a non-standard metaphysics of time. See Objections 5 and 6 below.

[6] What if Tim allows for the possibility of young-gramps' regeneration? Plausibly Tim intends to kill young-gramps in a permanent manner. For simplicity, I'll take killing to imply permanently-killing. In any case, Tim aims to do something that will make the future different from what he believes it was.

[7] I take causal relata to be states of affairs, although, for this paper, nothing significant hangs on this choice; I'll also talk loosely of other relata.

evidence of that state: symptoms are evidence of disease, fossils are evidence of dinosaurs. Evidence of this form has a dual character: it picks out both what is reliably correlated by physical means, as well as what we are generally responsible for forming empirical beliefs on the basis of. Evidence, as the terms is used here, isn't *factive*. Likewise, Tim's having a memory does not imply the memory is veridical.

Albert (2000) gives a related account of 'records'. Roughly, a local state A at time t_1, is a record of another local state B at an earlier or later time t_2, when they are correlated with one another, given chances derived from statistical mechanical laws (including a constraint on the initial state of the universe). Albert's notion, however, builds in a further condition: records must allow one to infer in ways that go beyond what one could infer to using laws and a probability distribution alone (without a constraint on the initial state). Because of this condition, Albert can explain a temporal asymmetry of records: why, in the actual world, we can have records of the past and not the future. But, for considering how physical states can constrain belief, it is useful to have a more general notion of evidence that includes both records, as well as local states that allow us to infer *merely* on the basis of laws and a probability distribution. Physical chances, including those derived from statistical-mechanical laws, can play this role. A counts as evidence of B just in case they are correlated given physical chances.

Records (explicated in Albert's terms) remain, nevertheless, a particularly important *type* of evidence, particularly in cases of time travel. A derived feature of records is that they allow one to infer to the recorded state of an object without having to know what happens to the object between now and then. For example, a photograph allows you to infer to what your grandmother looked like 20 years ago, without having to know what she looks like now, or at times in between. Albert uses statistical mechanics to explain why we can have records of the past and not the future in the actual world. But temporal asymmetry is not built into records at the conceptual level. It is conceptually possible to have records of the future—indeed, this is precisely what can happen in cases of backwards time travel.

I've also spoken of what evidence an agent *has*. An agent has evidence in virtue of having appropriate access to it. Tina might have evidence of the current state of the weather, for example, in virtue of looking out the window. Whether Tina *has* evidence depends on what observations she makes, and what kind of mental state she is in. What Tina's evidence is evidence is *of*, however, does not depend directly on her beliefs or other subjective states. If Tina looks out the window and sees storm clouds, she has evidence of later rain, even if she doesn't know about the correlation between dark clouds and later rain. Tina's evidence is of later rain because of the physical correlation between what she observes and later rain.

Let's return to time-travelling Tim. Take it that an agent is *evidentially rational* if their beliefs conform to their evidence. Assume that Tim is evidentially rational. Then he will be certain he won't succeed at killing young-gramps. His evidence now settles young-gramps' survival. Then, by *Ignorance*, Tim can't reasonably deliberate on killing young-gramps. Self-prediction rules out reasonable deliberation. In Tim's case, there is a tension between both following evidential norms, and reasonably deliberating. Assume further that Tim is not *deliberatively irrational*—

he does not engage in deliberation in contravention with *Ignorance*. Then Tim *won't* deliberate on killing his grandfather. Lewis, Horwich and Sider are right that merely the fact that Tim fails at his attempt to kill his grandfather doesn't prove he was unable to. But there is an important sense in which Tim's freedom as an agent *is* compromised. Tim's can't reasonably deliberate on killing young-gramps, certain he'll fail.

Tim may still be able to reasonably deliberate on other things—such as whether to *meet* young-gramps, whether he *ought* to kill young-gramps, the best way of doing so, or what would happen, were he to try. He may even be able to reasonably deliberate on *trying* to kill young-gramps, where trying consists in doing particular activities that would normally be expected to lead to killing, such as pulling the trigger of a loaded gun.[8] But Tim still can't reasonably deliberate on killing young-gramps. Moreover, there are other time travellers, such as those who encounter their younger selves, who would have even more significant limits on their deliberative freedom. If travellers retain memories and beliefs of having encountered their younger selves, for example, they may know precisely what they'll do before they do it.

Might Tim nevertheless still be *able* to kill young-gramps? My conclusion at this stage is explicitly limited to what Tim can reasonably deliberate on—although I will make broader claims about Tim's abilities in Sect. 6. One might think that a broader conclusion can quickly be drawn because of the link between intention and intentional action. According to a 'Simple View', discussed by Bratman (1984), an agent who intentionally ø's must have an intention to ø. So perhaps because Tim can't intend to kill young-gramps (by a related ignorance condition on intention), he can't intentionally kill him either. But even if the Simple View is correct, it has no direct implications for what one can do *simpliciter*. There are things I can do without doing so intentionally, such as alert a burglar to my presence. Even if one thinks that actions must be intentional under *some* description (Davidson 1980, essay 3), Tim's being unable to *intentionally* kill young-gramps doesn't yet imply that he is *unable* to kill young-gramps; Tim may be able to act intentionally under some other description. There is no straightforward inference from limits on intention to limits on action. This is why Rennick's (2015) conclusion in a similar case is limited to Tim's ability to *murder* young-gramps: murder, unlike most actions, explicitly requires intention.

3 Objections

Objection 1 Aren't you making the same error as a fatalist by assuming that facts about the future are fixed and smuggling them in by appealing to Tim's *knowledge* of the future?

---

[8] There are other senses of trying which Tim would be unreasonable to engage in (see Objection 8).

No. I didn't talk about knowledge, or its being a *fact* that Tim won't kill young-gramps. I appealed to Tim's *certainty* and *evidence* that he won't. Nor did I use a fatalist inference from Tim will fail to Tim must fail.

Objection 2   Surely 'certainty' isn't enough—agents can be certain for all sorts of reasons.

Agreed. Tim's case wouldn't be so interesting if Tim certainty in young-gramps' survival was mere delusion. But, given the causal and evidential structure of the case, Tim has *evidence* that he's a time traveller and that this is his young grandfather. Tim's memories and other physical evidence mean he can be justifiably certain about future events.

Objection 3   Couldn't Tim simply recover his ability to deliberate by taking an amnesia pill?

He could. But Tim's case is still of interest. In the actual world, we don't have records of our future behaviour. When we have evidence of our future behaviour, it goes by way of evidence about our present states—our decisions, character, and so forth. In cases where we're ignorant of our decisions and intentions, we typically don't have evidence of our future acts. So we can reasonably deliberate—even while remaining evidentially rational (and without taking an amnesia pill). But backwards time travel breaks these systematic conditions of ignorance, allowing agents to be justifiably certain of their future behaviour independently of their decisions. Even if time travellers can't always self-predict, the structure of these cases allows for this possibility.

Objection 4   Tim could never have sufficient evidence to justify certainty in young-gramps' survival. Perhaps he's been misled.

Recall that the certainty involved is 'practical certainty', which is relative to the practical and theoretical contexts relevant to the agents. Practical certainty can be justified, even if the possibility of error remains. Moreover, there is no principled limit to the amount of evidence Tim could acquire.

Objection 5   What if Tim is uncertain of the metaphysics of time, and countenances the possibility of branching universes, or multiple temporal dimensions? He may then be justifiably uncertain about whether he'll succeed.

Time travel scenarios make the possibility of alternative metaphysical theories of time more salient. So perhaps we wouldn't fault Tim for failing to follow his reliable evidence. But the epistemic possibility of alternative metaphysical theories doesn't undermine the reliability of Tim's evidence in his world. If Tim fails to self-predict, he is still not being responsive to the nomic and causal structure of his world. Moreover, time travellers may have many opportunities to test their reliability of their evidence. If so, it becomes less plausible that they could remain justifiably uncertain.

Objection 6 Agents wouldn't all of sudden be unable to deliberate. Instead, they'd fall prey to the 'second-time-around fallacy' (Smith 1997), and believe they can remake the future.

Perhaps. But it's still the case that Tim can *only* reasonably deliberate by failing to follow his reliable evidence. Furthermore, it's plausible that agents *wouldn't* continue to fall prey to this fallacy if they repeatedly tried to 'change' the future and failed to. We have no direct experience of time travel, and classic cases like Tim's involve only one-off attempts to change the future. Agents with more experience of time travel may find that their behaviour catches up with what it's rational for them to do.[9]

Objection 7 Given the rewards of killing young-gramps, Tim's deliberation remains reasonable.

Since practical certainty is relative to context, time travellers may sometimes reasonably deliberate on future states they have evidence of. But this won't always be the case. Say Tim has evidence that settles the results of his grandmother's early attempts at knitting. What would he gain by sabotaging these? There are cases where Tim's certainty of future states would be justified.

Objection 8 Tim would learn to make better decisions by deliberating, so his deliberation remains reasonable.

Just because an activity has benefits does not imply that it is reasonable. Trying to square the circle may be good practice at geometry, even if one knows that doing so is impossible. This doesn't make it reasonable. Moreover, Tim may be able to learn to make better decisions in ways that are compatible with satisfying *Ignorance*. He can think about what he ought to do, what options are available to him, and what his reasons are.

Objection 9 Isn't the problem really about the unreasonableness of intending or deciding when you're certain you'll fail (Smith 1997; Rennick 2015)?

It's not just certainty of failure that rules out reasonable deliberation, but certainty of intended result. Consider the following case. Tam travels back in time, intent on killing her young grandfather. After tracking him down, she discovers he's a sickly youth, not expected to survive the next winter. Tender-hearted Tam reneges on her previous intention, and deliberates about whether he is to survive. She decides he is, buys medicine, and tends at his bedside. Tam's deliberation is also unreasonable.[10]

---

[9] Such agents might continue to take themselves to act, even when they no longer deliberate, or might take themselves to become mere 'instruments of the world'. Smith (2005) explores other attitudes the time traveller might adopt.

[10] A similar result can be reached by requiring deliberation to be between incompatible alternatives: certainty of intended result rules out other incompatible options as available.

Objection 10   Isn't the problem rather about self-defeating causal loops: the fact that Tim's succeeding would produce an inconsistent set of events, where he both lives, and doesn't?

Self-defeating loops do introduce special puzzles. Tim's presence in the past implies his failure, given other assumptions (Vihvelin 1996). I'll discuss some of the connections between deliberation and self-defeating loops in Sect. 6. It's also true that features of the causal loop play a role in Tim's case. What allows Tim to have records of young-gramps' survival is partly the fact that future states *cause* past states. But a time traveller's freedom is constrained even in cases that don't involve causal loops.[11] Say Tim is justifiably certain that beside young-gramps is young-Maurice, who grows up to be gramps' business partner. Tim takes aim at young-Maurice. He intends to kill young-Maurice, not because of anything Maurice does in the future, but simply because Tim takes a dislike to him in the past. Tim's presence or activities in the past don't logically or nomically entail Maurice's survival. But Tim is equally unreasonable to deliberate on killing young-Maurice.

4 The open future

I've argued that time travellers like Tim can't reasonably deliberate on parts of the future they would usually control when their beliefs conform to their evidence. In the remaining sections, I'll draw out the consequences of this constraint. To begin, *Ignorance* can help explain why, in the actual world, we conceive of the future as open. Intuitively, we think of the future as unfixed and unsettled. Particularly when we deliberate, it seems that the future really could go one way or another. This appearance of openness is likely to be a multifaceted phenomenon, involving both phenomenological and belief-like aspects. *Ignorance* can help explain an aspect of it.

   An aspect of the future's apparent openness is that it appears open to control. We conceive of the future, but not the past, as something that is in principle up to us. While there are future states we don't take to be under our control, because they are too small, too large, or too far away, the future seems generically open to choice in a way the past is not. The past is fixed, and not even potentially under our control.

   One way of explaining this asymmetry is to appeal to causal asymmetry: the fact that causes come before their effects. The ignorance condition allows for a further, complementary, explanation. Begin with the fact that there are records of the past, but not the future (Albert 2000, chap. 6). There are local states of affairs now, such as the photograph of my grandmother, that allow one to reliably infer to the past states of objects, without knowing what happens to the object at times in between. Because there are no records of the future, we can only come to know future states by inferring from current (or past states) of objects and their surrounds, and

---

[11] If the causal relata are fine-grained, time travel without causal loops may not be possible. Even so, the time traveller's freedom would not be affected in virtue of the potentially self-defeating nature of the casual loop, but in virtue of its evidential features.

considering how they will likely evolve in the future. I may predict how my grandmother will look in 20 years time, but only by inferring from what she looks like now (or in the past), and taking into account how people age, and what is likely to befall her.

This asymmetry of records has consequences for self-prediction. When predicting our future behaviour, the best we can do is infer from what we know of our present or past states, such as our general character, intentions, and our past behaviour. This method doesn't typically allow us to predict much of our future actions, independently of making or inferring from our decisions. So *Ignorance* can be satisfied regarding much of our future behaviour, even when we follow our evidence. We can therefore reasonably deliberate on the future, while remaining evidentially rational. We can't similarly deliberate about recorded states of the past. It is asymmetries of evidence in the actual world, combined with evidential rationality, and *Ignorance*, that explain why the future appears open to our control in principle. However, if we had reliable access to the future that worked like memories, and gave us knowledge of our future behaviour independently of our decisions now, our freedom to reasonably deliberate on the future would be curtailed. This is precisely what can happen in cases of backwards time travel.

Does this mean that when we don't have records of past states, we can reasonably deliberate on them? Plausibly there are additional requirements on reasonable deliberation. Reasonable deliberation requires taking our decision to evidentially settle the option under deliberation. Based on defenses of evidential decision theory (Price 1991), I've argued elsewhere (2017, 2019) that this condition won't be satisfied for past states in the actual world, for agents who follow their evidence. But the condition can be satisfied for future states. So agents can reasonably deliberate on some future states (while following their evidence), but no past states.

Can this kind of account explain why the *whole* future appears open to control, while *none* of the past does? One possibility is that we generalise from *some* future states being under our control to *all* future states being *potentially* under our control (and from *no* past states being under our control to no past states being *potentially* under our control).[12] Another possibility is that we take all past states (but no future states) to be *knowable* by records, and so knowable independently of our decisions now. This might lead us to believe that past states aren't *potential* objects of reasonable deliberation, while future states are. Note that none of these generalising moves are strictly licensed by *Ignorance*, which merely concerns what a deliberating agent can reasonably deliberate on, given her beliefs—not what she can potentially deliberate on.

Various other asymmetries no doubt contribute to why the future appears potentially under our control. These include the way memories and experiences accumulate (Mellor 1998, pp. 122–123; Ismael 2011), and how we infer to the past and future (Albert 2000, Chap. 6), independently of how this affects deliberation. It's also plausible that an intuitive metaphysical conception of time feeds back into why we conceive of the future as evidentially open, and not something we could, in

---

[12] Generalising moves of a similar form are made by Prosser (2016, Chap. 7) and Ismael (2007, 2011).

principle, come to know. Ultimately why we conceive of the future as open is a matter to be settled by empirical investigation. What I point to here is one way in which constraints on reasonable deliberation can play a role.

5 More time travel, less bananas

A second important consequence of the constraint on deliberation concerns the possibility of time travel. Cases like Tim's have been used to argue against the possibility of time travel and backwards causation. In this section I show how *Ignorance*, combined with additional assumptions, undercuts these arguments.

Horwich (1987, pp. 119–121) argues against the epistemic possibility of time travel to the recent past as follows. If such time travel were to occur, there would be lots of attempts to bring about self-defeating causal loops, such as Tim's attempts to kill young-gramps. These attempts would lead to long strings of apparently-miraculous coincidences—Tim slipping on a banana-peel, the gun jamming, uncharacteristic changes of heart. We have good reason to think such long strings of coincidences won't take place. So we have good reason to believe time travel to the recent past will not take place (or at least not often).

Smith (1997, pp. 379–382) rejects this conclusion, arguing that unlikely outputs only result from unlikely inputs. A time traveller like Tim trying to bring about states he is certain won't happen (such as killing young-gramps) is an unlikely state of affairs. So time travel doesn't imply unlikely states follow from likely ones, and we have no reason to think time travel itself is unlikely. *Ignorance* provides a principled route to a similar conclusion. While the condition does not settle how likely various psychological states are, it holds that agents are unreasonable when they deliberate on options they are certain of. If agents' behaviour in the long term catches up to what it is rational for them to do, then agents won't continually decide to bring about states they are certain won't happen. So we shouldn't expect long strings of apparently-miraculous coincidences. Strings of coincidences will be, at most, short-lived intermediaries.[13]

Not all attempts to instantiate self-defeating causal loops involve agents. A more general response to cases involving self-defeat is to consider how the probable behaviour of a system changes when causal loops are introduced (Wheeler and Feynman 1949; Arntzenius and Maudlin 2013). Our intuitions about likely behaviour become unreliable when we consider systems containing causal loops. What *Ignorance* helps reveal is how our intuitions about our *own* likely behaviour can also be misleading. While we may have thought, intuitively, that attempts to remake the future would be the inevitable result of human psychology, we will have to be evidentially *irrational* if we're to reasonably engage in such attempts.

Ignorance also undercuts so-called 'bilking' arguments against the possibility of backwards causation (Black 1956). These arguments take the following form. Say I become convinced that my train-waiting ritual, *A*, causes the train to leave the

---

[13] For responses to objections involving curiosity, see Smith (1997, pp. 383–385).

previous station on time, B. A putatively causes B, even though A comes *after B*. But it seems I can break any apparent correlation between the two as follows. I call to check whether the trains left the previous station on time. If the train *has* left on time I *don't* do the ritual. If the train *hasn't* left on time, I *do* the ritual. I thereby 'bilk' the correlation between A and B. Because the correlation is bilkable, A can't be a cause of B.

Michael Dummett (1964) and Huw Price (1996, pp. 171–174) use similar reasoning to argue that backwards causation requires principled ignorance of the past. A can only cause B if B is unknowable, so that the correlation is not bilkable.[14] *Ignorance* shows that no such restriction is required. If an agent has evidence of B, and A is reliably correlated with B, she has evidence that settles A. So she can't reasonably deliberate on A if she follows her evidence. Bilking the correlation requires a long series of trials in which an agent *decides* to bring about not-A when she is certain of B. If her deciding is governed by to what she can reasonably deliberate on while following her evidence, she won't bilk the correlation. So backwards causal relations won't be bilked, even if there is no principled ignorance (or spooky inabilities) on the part of the agent.[15]

Similar responses apply in cases that involve records of the future, such as Goldman's (1970) 'Book of Life'—a book that recounts someone's complete life in detail. One might argue (*contra* Goldman) that such a book is impossible, or unlikely, because, if the main character (call her Osma) were to read the book, she would attempt to thwart its predictions. Either she succeeds, implying the book isn't a Book of Life, or a long series of apparently miraculous coincidences is required to prevent her. If we think such a series is unlikely, it seems we should conclude there aren't likely to be any Books of Life, or at least not any we might encounter. But, if Osma deliberates in accordance with *Ignorance* and follows her evidence, she won't deliberate on thwarting the book's predictions. If she decides and acts via reasonable deliberation, she won't bilk the predictions. Nor will a long series of apparently-miraculous coincidences be required to prevent her succeeding.

6 Counterfactuals and abilities

I've argued that time travellers like Tim can't reasonably deliberate on states they would ordinarily control in the actual world (such as killing young-gramps) given that they're certain they'll fail. But it might seem that this result is of limited interest, and misses the point. Who cares whether Tim can reasonably deliberate? Either he *can* kill grandfather or he can't. And never mind agents. Shouldn't we be concerned about the abilities of objects more generally? Surely questions of how to evaluate abilities are more tractable when messy agents are left out of the picture. In this final section, I explain how rational constraints on deliberation *are* relevant to

---

[14] More carefully, B can't be known in a way that would leave the correlation between A and B intact.

[15] One might still argue that an agent *can* bilk the correlation, even if she doesn't. But then one would need to hear more about why this sense of 'can' is relevant to the correlation being causal, rather than one more closely tied to what an agent can reasonably deliberate on.

the evaluation of abilities, particularly in cases of time travel. I'll use these constraints to motivate evidential and temporally neutral approaches to evaluating *counterfactuals*, from which abilities are then derived. The methods I'll suggest have the unusual feature of breaking a straightforward connection between abilities, counterfactuals and causation in cases of time travel. As I'll argue, while this is an odd result, it contributes to explaining part of what's puzzling about causal loops. Note that while I'll sometimes speak as though there is a single right way of evaluating counterfactuals and abilities, my view is that the evaluation of both is context-sensitive. I return to this point below.

One reason that rational constraints may seem beside the point is that it may seem Tim straightforwardly *can* kill his grandfather. Tim's abilities are determined by his local causal powers (and perhaps his present local surrounds). Tim's abilities are therefore unaffected by time travel: Tim can normally kill a man he confronts with a loaded gun, so Tim *can* kill young-gramps. But this reasoning is too quick. Even if we take the local causal structure of the world as primitive, one still needs to work out abilities that are manifest at some distance. In Tim's case, we're unlikely to treat Tim's ability to kill young-gramps as a primitive, since the killing happens causally downstream of Tim's direct actions, such as his bodily movements.

One standard method for determining Tim's abilities is to relate then to counterfactuals about what would happen, were he to decide, intend or try.[16] Tim can kill his grandfather *if and only if*, were he to decide, intend, or try to, he would or might succeed.[17] I'll assume this method in what follows. We then need to evaluate these counterfactuals. It might seem we can do so in a relatively straightforward manner, by holding the past of the antecedent fixed and not holding any future states fixed.[18] Aspects of the future may counterfactually depend on Tim's actions now, but no aspects of the past. Because grandfather's survival is in the future of Tim's actions, and we would expect a (causal) correlation between Tim's actions and young-gramps' survival, young-gramps' survival counterfactually depends on Tim's actions. So Tim can kill young-gramps.

But things are not so straightforward. Cases of backwards time travel put pressure on temporally asymmetric methods of evaluating counterfactuals. Backwards time travel requires backwards causation. If an asymmetry of counterfactual dependence explains the asymmetry of causation in the actual world (Lewis 1979), then cases of backwards time travel *require* backwards counterfactual dependence. So we can't evaluate counterfactuals by holding the past fixed. Similarly, if counterfactual asymmetry is merely related to causal asymmetry (perhaps causal asymmetry explains counterfactual asymmetry, or they are both explained by some third asymmetry), then we would expect cases of backwards time travel to lead to backwards counterfactual dependence. Tim arriving in the past, for example, counterfactually depends on his entering the time machine.

---

[16] If counterfactuals are context-sensitive, we may need to specify that these are the kinds of counterfactuals relevant for decision-making and abilities.

[17] Arguably, this is a necessary but not sufficient condition, since a causal condition may be required as well. My arguments won't depend on this detail, or others concerning the condition's content.

[18] Perhaps ignoring a 'transition period' just prior to the antecedent (Lewis 1979).

Cases of backwards time travel also put pressure on methods that hold *no* future states fixed. In cases of backwards time travel, future states have features that usually only hold of past states in the actual world. Firstly, states in the future *cause* states in the present. Tim's presence now is caused by his future departure. If we use a method of evaluating counterfactuals that holds the 'causal past' of aspects of the present fixed, then we will be led to hold aspects of the temporal future fixed. For example, we might hold Tim's future departure fixed when evaluating counterfactuals concerning Tim's actions in the present.[19] Secondly, in cases of backwards time travel, there can be *records* of the future. There can be local states that allow one to reliably infer to Tim's future behaviour, independently of knowing his current state. If we use a method of evaluating counterfactuals that holds states recorded in the present outside the area of the antecedent fixed (Albert 2000, Chap. 6; Loewer 2007), we will be led to hold future states fixed. For further details on these arguments, see Fernandes (2018).

Rather than using a temporally asymmetric method, can we evaluate counterfactuals and abilities in more temporally *neutral* terms? Rational constraints on deliberation give us some guidance. Here's a suggestion. A plausible thought is that counterfactuals and ability facts concerning Tim should be relevant to Tim. Given Tim has reliable evidence of future states, independently of the decisions he makes now, these future states should be held fixed when considering what would happen, were Tim to decide or act in different ways. Because Tim has reliable evidence of young-gramps' survival, even while he is justifiably uncertain of his decision and actions, young-gramps' survival is held fixed when determining what would happen were Tim to decide to kill young-gramps, or try to. If Tim's abilities are determined by counterfactuals evaluated in these terms, Tim *can't* kill young-gramps. It is not the case that, were Tim to decide, intend or try to kill young-gramps, he would or might succeed. By this method, Tim's abilities aren't determined merely by his local features and surrounds. States at more distant times, such as young-gramps' future survival, are relevant to what Tim can do.

As observers of Tim, we may be tempted to still evaluate counterfactuals (and abilities) by not holding *any* future states fixed, and so not holding young-gramps' survival fixed (Lewis 1976; Sider 2002; Ismael 2003). If we do, Tim presumably *can* kill young-gramps—in which case Tim turns out not to be the time-travelling grandson of young-gramps. But I suspect that some of our reluctance to hold future states fixed in time travel cases arises from simply *presuming* a temporally asymmetric method of evaluating counterfactuals. A temporally asymmetric method works reasonably well in the actual world, where we never have records of the future, and so lack evidential reasons for holding future states fixed. But in Tim's world there are records of the future. Agents like Tim can reliably self-predict the future, independently of the decisions they make now. Insofar as evaluations of abilities and counterfactuals are to be relevant to these deliberating agents, they

---

[19] For a related argument, see Kiourti (2008, pp. 349–350). Kiourti argues that future events are relevant to the time traveller's ability to kill her younger-self because they're in her 'personal' past. The argument I give here is more general: it does not rely on demarcating the time-traveller's personal past, or on facts concerning the identity of the time traveller.

need to be sensitive to the kind of evidence that these agents have. If agents have reliable evidence of the future, parts of the future need to be held fixed when evaluating counterfactuals and abilities concerning these agents.

Does such an approach make counterfactuals unreasonably subjective? Note that which counterfactuals are true does not depend on what agents happen to believe or desire. Recall, Tim's evidence depends on what the states accessible to him now are reliably correlated with, not what he happens to believe or want. But one could adopt a more objective variant of the above approach. For example, one could hold fixed states settled by the evidence that agents *in Tim's situation* could be expected to have (where Tim's situation is specified in causal terms). Or one could appeal to the evidence that agents *surrounding* Tim would be expected to have. Alternatively, one could take evaluations of counterfactuals (and abilities) to be relative to particular agents, or evidence sets. Under any of these approaches, it is not straightforwardly true that Tim *can* kill young-gramps. Either it is false (if we consider Tim's evidence, or evidence of agents in Tim's situation), false relative to some evidence sets, or depends on Tim's surrounds.

My own view is that counterfactuals and abilities are context-sensitive. Depending on our needs, we may evaluate counterfactuals and abilities using any of the above approaches, including temporally asymmetric ones. My point is that there are reasonable contexts in which parts of the future are held fixed, and Tim cannot kill young-gramps. These are likely to be contexts that focus on the needs of the deliberating agent.

Keeping context-sensitivity in mind, here's *a* temporally-neutral method of evaluating counterfactuals. This method concerns counterfactuals where (a) the antecedent is an agent's decision, and (b) her evidence does not settle her decision. To evaluate such counterfactuals, hold fixed what the deliberating agent's evidence settles, independently of her decision. To evaluate counterfactuals concerning Tim's decision to pull the trigger, for example, we hold fixed young-gramps' survival. Why? Because Tim's evidence settles young-gramps' survival, independently of his decision. To work out what counterfactuals are true, we then consider what Tim's decision will be evidence for. If Tim's decision to pull the trigger evidentially settles the gun firing, for example, then the gun's firing counterfactually depends on Tim's decision to pull the trigger. But young-gramps' *survival* does not counterfactually depend on Tim's decision to pull the trigger, since Tim's evidence settles grandfather's survival independently of his decision.

This method extends to antecedents that are actions, provided (c) the agent only deliberates or decides on states of affairs her evidence does not settle. Say Tim is such an agent, and deliberates on whether to pull the trigger. As before, states that the deliberating agent's evidence settles are held fixed. States depend counterfactually on the agent's action, just in case they are evidentially settled by the agent's action, when the action is settled by decision. For example, the firing of the gun counterfactually depends on Tim's pulling the trigger just in case Tim's pulling the trigger, when he decides to, evidentially settles the gun firing.

We can derive Tim's abilities from counterfactuals evaluated in the above terms. Assume that similar methods also apply to intentions and trying—they can be treated as actions or decisions for the purposes of evaluating counterfactuals. Then,

using these methods, Tim can't kill young-gramps. It is false that, were Tim to decide, intend or try to kill young-gramps, he would or might succeed.

Much more would need to be said to extend these methods to a full range of counterfactual antecedents. My suspicion is that the evaluation of counterfactuals become highly context-sensitive once we move beyond contexts relevant to deliberating agents. One suggestion is that we sometimes evaluate counterfactuals by finding a plausible decision point where an agent could have reasonably deliberated and decided in one way or another (even if no deliberating agent was in fact there)—see Fernandes (2017, pp. 13–15) for further discussion.

In addition to being temporally neutral, the above evidential methods of evaluating counterfactuals can explain part of what's puzzling about causal loops. Causal loops allow agents to have evidential access (independently of their decisions) to states that they would be expected to control in the actual world. Tim can have evidence of young-gramps' survival, independently of his decision—even though, in ordinary cases, we'd expect young-gramps' survival to be up to him. What's particularly odd is that, given this is a causal loop, young-gramps' survival still *causally* depends on Tim's actions—even though, under the above evidential approaches to evaluating counterfactuals, Tim *isn't able* to kill young-gramps. Tim isn't able to do things that causally depend on his actions.

Has something gone wrong? Not necessarily. What's happening in Tim's case is that a straightforward connection between counterfactuals, abilities and causation *isn't* being maintained for agents who can't reasonably deliberate (given they follow their evidence). If one takes causation to be paradigmatically relevant for deliberating agents (Price 1991; Pearl 2000; Woodward 2003; Ismael 2007; Blanchard 2014; Fernandes 2017), it's no longer surprising that causal information isn't relevant for agents who systematically *can't* properly deliberate on given states. Causal loops are then puzzling because they systematically undermine the possibility of deliberatively (and evidentially) rational agents making use of the causal relations of which they're composed. If we require evaluations of abilities to still be relevant to deliberating agents in such circumstances, as I think we should, a straightforward link between causation, counterfactuals and abilities is broken.

7 Conclusion

Time travelling agents don't always retain their ordinary freedoms and abilities when they travel back in time. If Tim has evidence of young-gramps' survival, independently of the decision he makes now, he should self-predict, and be certain he won't succeed in killing young-gramps. If he does, then, given an ignorance condition on deliberation, Tim can't reasonably deliberate on killing young-gramps. He can't reasonably deliberate on states we would ordinarily expect him to control.

The ignorance condition helps explain the apparent openness of the future in the actual world. We take the future to be open to deliberation, but not the past. This is partly explained by the fact that, in the actual world, we have records of the past, and not the future. Given *Ignorance*, we can't reasonably deliberate on the recorded past while following our evidence. But no such constraint affects our ability to

deliberate on the future. In situations that involve backwards causation, however, our ability to deliberate on the future is curtailed—affecting what we say about the possibility of time travel and backwards causation. Bilking and other arguments against the possibility of backwards causation rely on the idea that an agent's ability to deliberate is *not* affected by her evidence. But, if rational agents don't continually try to bring about states that their evidence settles won't happen, bilking and other arguments against time travel are undercut.

Finally, the ignorance constraint on deliberation motivates a temporally neutral and evidence-based method of evaluating counterfactuals and abilities. According to this method, counterfactuals and abilities concerning an agent's decisions and actions are evaluated by holding fixed states the agent has reliable evidence of, independently of her decision—whether these states lie in the past or future. While we might be tempted to evaluate counterfactuals by holding only past states fixed, this temptation reflects merely contingent temporal asymmetries in the actual world—that we have records of the past and not the future.

In these ways, rational constraints on deliberation turn out to matter to what may have seemed like distinctly metaphysical issues: the freedom of agents, the openness of the future, the possibility of backwards causation, and the evaluation of counterfactuals and abilities.

References

Albert, D. (2000). *Time and chance*. Cambridge, MA: Harvard University Press.

Arntzenius, F., & Maudlin, T. (2013). Time travel and modern physics. In Edward N. Zalta (Ed.) *The stanford encyclopedia of philosophy* (Winter 2013 Edition), http://plato.stanford.edu/archives/win2013/entries/time-travel-phys/.

Black, M. (1956). Why cannot an effect precede its cause? *Analysis, 16,* 49–58.

Blanchard, T. (2014). Causation in a Physical World. PhD Thesis. Rutgers University.

Bratman, M. (1984). Two faces of intention. *The Philosophical Review, 93*(3), 375–405.

Davidson, D. (1980). *Essays on actions and events*. Oxford: Clarendon Press.

Dennett, D. (1984). *Elbow room*. Cambridge, MA: MIT Press.

Dummett, M. (1964). Bringing about the past. *Philosophical Review, 73,* 338–359.

Fernandes, A. (2016a). Varieties of epistemic freedom. *Australasian Journal of Philosophy, 94*(4), 736–751.

Fernandes, A. (2016b). A deliberative account of causation: How the evidence of deliberating agents accounts for causation and its temporal direction. PhD Thesis. Columbia University.

Fernandes, A. (2017). A deliberative approach to causation. *Philosophy and Phenomenological Research, 95*(3), 686–708.

Fernandes, A. (2018). Time travel and counterfactual asymmetry. **(unpublished manuscript)**.

Fernandes, A. (2019). Time, flies, and why we can't control the past. In B. Loewer, E. Winsberg, B. Weslake (Eds.), *Time's arrows and the probability structure of the world*. Cambridge, MA: Harvard University Press (in press).

Ginet, C. (1962). Can the will be caused? *The Philosophical Review, 71*(1), 49–55.

Grice, H. P. (1971). Intention and uncertainty. *Proceedings of the British Academy, 57,* 263–279.

Hampshire, S., & Hart, H. L. A. (1958). Decision, intention and certainty. *Mind, 67*(265), 1–12.

Harman, G. (1976). Practical reasoning. *The Review of Metaphysics, 29*(3), 431–463.

Holton, R. (2006). The act of choice. *Philosophers' Imprint, 6*(3), 1–15.

Horwich, P. (1987). *Asymmetries in time*. Cambridge MA: MIT Press.

Ismael, J. (2003). Closed causal loops and the bilking argument. *Synthese, 136,* 305–320.

Ismael, J. (2007). Freedom, compulsion and causation. *Psyche, 13*(1), 1–11.

Ismael, J. (2011). Temporal experience. In C. Callender (Ed.), *The Oxford handbook of philosophy of time*. Oxford: Oxford University Press.

Kapitan, T. (1986). Deliberation and the presumption of open alternatives. *The Philosophical Quarterly, 36*(143), 230–251.

Kiourti, I. (2008). Killing baby suzy. *Philosophical Studies, 139*(3), 343–352.

Levi, I. (1980). *The enterprise of knowledge*. Cambridge, MA: MIT Press.

Levi, I. (1986). *Hard choices*. Cambridge: Cambridge University Press.

Lewis, D. (1976). The paradoxes of time travel. *American Philosophical Quarterly, 13,* 145–152.

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs, 13*(4), 455–476.

Loewer, B. (2007). Counterfactuals and the second law. In H. Price & R. Corry (Eds.), *Causation, physics, and the constitution of reality*. New York: Oxford University Press.

Mellor, D. Hugh. (1998). *Real time II*. London: Routledge.

Owens, D. (2008). Deliberation and the first person. In A. E. Hatzimoysis (Ed.), *Self-knowledge*. Oxford: Oxford University Press.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge: Cambridge University Press.

Price, H. (1991). Agency and probabilistic causality. *British Journal for the Philosophy of Science, 42,* 157–176.

Price, H. (1996). *Time's arrow & archimedes' point*. New York: Oxford University Press.

Prosser, S. (2016). *Experiencing time*. New York: Oxford University Press.

Rennick, S. (2015). Things mere mortals can do, but philosophers can't. *Analysis, 75*(1), 22–26.

Sider, T. (2002). Time travel, coincidences and counterfactuals. *Philosophical Studies, 110,* 115–138.

Smith, N. J. J. (1997). Bananas enough for time travel? *British Journal for the Philosophy of Science, 48,* 363–389.

Smith, N. J. J. (2005). Why would time travellers try to kill their younger selves? *Monist, 88*(3), 388–395.

Vihvelin, K. (1996). What time travelers cannot do. *Philosophical Studies, 81,* 315–330.

Wheeler, J., & Feynman, R. (1949). Classical electrodynamics in terms of direct interparticle action. *Reviews of Modern Physics, 21,* 425–434.

Woodward, J. (2003). *Making things happen*. New York: Oxford University Press.