Synthese (2024) 204:131
https://doi.org/10.1007/s11229-024-04740-7

ORIGINAL RESEARCH

The role of causal manipulability in the manifestation of time biases

Batoul Hodroj[1] · Andrew J. Latham[2] · Kristie Miller[1] · Rasmus Pedersen[1] · Danqi Wang[3]

Received: 5 January 2024 / Accepted: 3 August 2024 / Published online: 2 October 2024
© The Author(s) 2024

Abstract
We investigate the causal manipulability hypothesis, according to which what partly explains (a) why people tend to prefer negative events to be in their further future rather than their nearer future and positive events to be in their nearer future rather than their further future and (b) why people tend to prefer that negative events be located in their past not their future and that positive events be located in their future not their past, is that people tend to discount the value of events they are less able causally to manipulate. If people discount the value of events they are less able to causally manipulate, then since often temporally nearer future events are more causally manipulable than further ones, and since future events are manipulable whereas past ones aren't, this could explain both (a) and (b). In turn, if the causal manipulability hypothesis is correct, this might suggest that insofar as people's preferences are explained in this manner, those preferences are rationally permissible, since relative causal manipulability is normatively relevant. Thus, ascertaining whether the causal manipulability hypothesis is true may shed light on the normative status of such preferences. We investigate the cognitively mediated version of the causal manipulability hypothesis, according to which people's consciously held beliefs about the relative causal manipulability of events explains (a) and (b). Contrary to expectations, we found no evidence in favour of this view. We suggest that either relative causal manipulability plays an explanatory role, but one that is not cognitively mediated, or that it does not play any role, even if it is sometimes associated with other factors, such as probability, that may play a role in explaining some time biased preferences.

Keywords Future bias · Near bias · Experimental philosophy · Preferences · Time bias · Causal manipulation

1 Introduction

Suppose Freddie prefers to eat cake now rather than at the end of the day. Then Freddie is *apparently time biased.* He has a preference for where in time the cake consumption

Extended author information available on the last page of the article

occurs. There are various reasons he might have this preference. Freddie might believe that the cake will get progressively staler throughout the day, and hence that the cake later in the day will be less tasty than the cake now. Freddie might believe that someone else will eat the cake during the day if he does not eat it now, and so believe that the probability of *in fact eating* the cake now is higher than the probability of eating the cake later. If factors such as these are the sole reason for Freddie's preference, such that absent such factors he would have no preference about when he eats the cake (that is, holding all else equal) then Freddie is *merely apparently* time biased. Freddie's preferences are not sensitive to where in time the cake eating is located; they are sensitive to the properties of the cake (its staleness) and to the probability of eating the cake. By contrast, Freddie might be someone who continues to have a preference for where in time the cake eating occurs, holding fixed all such other facts. In that case, his preference is sensitive to the temporal location of the cake per se. If that is so, we will say that Freddie is time biased. Time biased preferences, then, are preferences that are sensitive to where in time events are represented as being located.[1]

An agent is *apparently near biased* if they tend to prefer positive events[2] to be located temporally near, and negative events[3] to be located temporally far, and they are near biased if they have this preference, all else being equal.[4] More carefully, we will say that an agent is apparently *prospectively* near biased if they tend to prefer positive events to be located in the near future rather than the further future, and to prefer negative events to be located in the further future rather than the near future. Agents are prospectively near biased if they have this preference all else being equal. Agents are apparently *retrospectively* near biased if they prefer that positive events are located in the near past rather than the further past, and negative events to be located in the further past rather than the nearer past. An agent is retrospectively near biased if they have this preference all else being equal.

An agent is *apparently future biased* if they tend to prefer that positive events are located in the future not the past, and that negative events are located in the past, not the future. An agent is *future biased* if, all else being equal, they have this preference. Since agents often know whether an event occurred in the past or not, future biased preferences are typically elicited by asking someone to imagine that there is some event, such as a painful operation, but that they do not, as of the present moment, remember whether the event already occurred or is still to occur. They are then asked which of these they prefer is the case. For example, here is Parfit (1984, p. 165) describing his preferences when it comes to the temporal locations of pains, in what we will call Parfit's Operation.

[1] Notice that preferences count as being time biased, as opposed to merely apparently time biased, if they are sensitive in this manner. It is not required that they are insensitive to other factors, such as probability, or value. Time biased preferences that are only sensitive to the temporal location of events and are not sensitive to any factors that *accompany* temporal location (such as probability, etc.), are known as *purely* time biased preferences, and it is very controversial whether anyone has such preferences.

[2] That is, events that the agent takes to have positive utility.

[3] That is, events the agent takes to have negative utility.

[4] In economics and psychology this is sometimes known as having a high time preference (as opposed to having a low time preference). For example, see Fredrick et al. (2002) and Lawless et al. (2013).

> I am in some hospital, to have some kind of surgery. Since this is completely safe and always successful, I have no fears about the effects. The surgery may be brief, or it may instead take a long time. Because I have to cooperate with the surgeon, I cannot have anaesthetics. I have had this surgery once before, and I can remember how painful it is. Under the new policy, because the operation is so painful, patients are now afterwards made to forget it. Some drug removes their memories of the last few hours.
> I have just woken up. I cannot remember going to sleep. I ask my nurse if it has been decided when my operation is to be, and how long it must take. She says that she knows the facts about both me and another patient, but that she cannot remember which facts apply to whom. She can tell me only that the following is true. I may be the patient who had his operation yesterday. In that case, my operation was the longest ever performed, lasting ten hours. I may instead be the patient who is to have a short operation later today. It is either true that I did suffer for ten hours, or true that I shall suffer for one hour.
> I ask the nurse to find out which is true. While she is away, it is clear to me which I prefer to be true. If I learn that the first is true, I shall be greatly relieved.

Parfit reports that he would strongly prefer that his operation be in the past, and he predicts that others in the same position would have the same preference. Even though the past operation if he had it, was much longer and contained more pain than the future operation. Parfit's preference is an instance of (apparent) future bias.

Research shows that people are *apparently* retrospectively near biased (Bickel et al., 2008; Yi et al., 2006; Greene et al., 2021a, 2021b),[5] apparently prospectively near biased (for a comprehensive sample of historical research on intertemporal choice see Loewenstein & Elster, 1992) and apparently future biased (Caruso et al., 2008; Greene et al., 2021a Latham et al., 2020; Latham et al., 2021a, 2021b, 2023). Evidence also shows that people are what we might call *strongly* near and future biased. That is, people will prefer less of a positive thing in the near past/future to more of a positive thing in the far past/future (strong near bias; Bickel et al., 2008; Yi et al., 2006; Critchfield & Kollins, 2001) and will prefer more of a negative thing when it is past compared to less of a negative thing when it is future, and less of a negative thing when it is future, compared to more when it is past (strong future bias; Greene et al., 2021a, 2021b). Finally, research shows that people are near biased with respect to a range of goods including money (Ainslie, 2005; Critchfield & Kollins, 2001) and that they are future biased with respect to a range of different kinds of experiences including visual, auditory, and taste sensations as well as pain/pleasure (Latham et al. 2023).

It remains an open question whether at least some of these studies also show that people are near biased and future biased because it is hard to know whether they have controlled for *all* the factors to which people's preferences might be sensitive, other

[5] Bickel et al. (2008) and Yi et al.(2006) found that people display the same patterns of apparent retrospective near bias, at population level, as they do apparent prospective near bias. Likewise, Greene et al. (2021a, 2021b) also found the same percentage of people reporting apparently retrospectively near biased preferences as prospectively near biased ones. Interestingly, though, they also found that there was no association between individuals being prospectively near biased and being retrospectively near biased.

than just temporal location, so it is hard to know whether people would retain the preference all else being equal (since it is hard to make sure all else is equal).

Determining which factors explain our having apparently time biased preferences is important in evaluating the rationality of those preferences. Many philosophers have argued that preferences that are sensitive to mere temporal location are sensitive to something that is normatively irrelevant, and in this respect, those preferences are not rationally permissible. For instance, Sidgwick, (1874, pp. 381) writes, "The mere difference of priority and posteriority in time is not a reasonable ground for having more regard to the consciousness of one moment than to that of another. The form in which it practically presents itself to most men is 'that a smaller present good is not to be preferred to a greater future good' (allowing for difference of certainty)." Rawls (1971, p. 293) argues that "A present or near future advantage may be counted more heavily on account of its greater certainty or probability, and we should take into consideration how our situation and capacity for particular enjoyments will change. But none of these things justifies our preferring a lesser present to a greater future good simply because of its nearer temporal position.[6]

If, however, it turns out that apparently time biased preferences are *merely apparently* time biased, because they are in fact explained by people's preferences *only* being sensitive to factors other than temporal location, it might turn out that some or all of these other factors are normatively relevant. For instance, the staleness of the cake that Freddie is considering is surely normatively relevant to his preference for where the cake is located in time, and likewise is the relative probability of him getting to eat the cake.

One possibility is that one factor that partially explains *both* apparent prospective near bias and apparent future bias is the relative causal manipulability of events. We will say that an agent can *causally manipulate* an event if they can cause the event to occur, cause the event to have certain properties rather than others,[7] or cause the event to fail to occur. In what follows we make two assumptions. First, it is *possible* to causally intervene on past events. That is, backwards causation is possible. An agent can causally manipulate a past event if, by acting in the present, they can bring it about that that event occurs, or that it does not occur. *Mutatis mutandis* for future events. Second, we assume that backwards causation is *not actual*, or at the very least, there is no backwards causation of the sort that would enable agents like us to bring it about that the past in some way we prefer.[8] We take these to be very plausible assumptions. Given these assumptions, it follows that for actual agents, past events are not causally manipulable, whereas (at least some) future events are. So, the general idea of the manipulability hypothesis is that because actually, past events are *not* causally manipulable, actual agents devalue past events compared to future ones

[6] Greene & Sullivan (2015, Sect. 2), Lewis (1986), Nagel (1970), Broome (1991), and Brink (2010) also make the same point.

[7] Here we assume a fairly coarse-grained individuation of events.

[8] Perhaps, for instance, people like Price (1994, 1996) are right and the correct explanation of Einstein–Podolsky–Rosen (EPR) correlations in quantum mechanics is quantum backwards causation. However, even if that is right, even Price does not think that human agents can use information to causally intervene on the past.

and therefore show apparent future bias, and that because far future events are less causally manipulable than nearer ones, people tend to value nearer events more highly, and thus show apparent near bias.

Latham et al. (2023) recently suggested something along these lines. Their suggestion is in line with that of other philosophers who have previously argued that we care more about the future than the past because the past is *practically irrelevant* to us (Horwich, 1987; Kauppinen, 2018; Latham et al., 2020; Maclaurin & Dyke, 2002; Parfit, 1984; Suhler & Callender, 2012). The idea is that we attach less evaluative weight to past events because there is nothing we can do to affect the past, which means that past events cannot count for, or against, present choices in the way that potential future events do. This idea was perhaps first suggested by Hume, who writes that the greater effect of future events than past events on the will 'is easily accounted for[:] As none of our actions can alter the past, 'tis not strange it shou'd never determine the will' (Hume, 1738, sect. 2.3.7.6). More recently, Kauppinen (2018) has argued that apparently future biased preferences are rationally justified (and, he seems to think, explained) by the fact that they have no effect on our choices.[9] Our inability to affect the past also underlies a popular evolutionary explanation for apparent future bias suggested by Parfit (1984, p. 186) and Horwich (1987, pp. 194–196) and developed by Maclaurin and Dyke (2002) and Suhler and Callender (2012).

At this point, it is useful to stop and consider what is meant by 'preference' in the context of discussion of apparent time biases, and in particular, apparent future bias. Preferences provide information about the *relative* attractiveness of different states of affairs. One common view about preferences is that preferences *just are* certain patterns of observed behaviour. So-called revealed preference theory (Samuelson, 1938, 1948, 1950) is often interpreted in this manner. By contrast, *mentalist* views of preferences take preferences to be mental states; in particular, evaluative attitudes. Philosophers concerned with time biases typically take preferences to be mental states. After all, consider again the case of Parfit's Operation. Parfit claims to prefer the 10-h past operation to the future 1-h operation. But Parfit cannot behave in such a way as to bring about the 10-h past operation over the 1-h future operation. So, if having that preference just is behaving in that way, then it's not the case that Parfit has this preference. Since no one can *actually* behave in such a way as to bring about some event in the past; it is not possible to have a pattern of behaviour that would constitute there being apparently future biased preferences. Thus, to make sense of apparent future bias, we, like those before us, assume a mentalist view of preferences according to which preferences are evaluative mental states.Then sometimes (perhaps often) preferences are inputs into decisions, as when we choose to perform one action over another. But sometimes preferences, so understood, are not inputs into decision because there are no actions that can be taken to bring about the object of the preference. In these latter cases, however, individuals still *have* the preference, and can still *report* the preference, even if they cannot act on it. This is the case when it comes to apparently future biased preferences. In such cases, we assume that individuals can *report* how they *would* act

[9] Kauppinen (2018) does not claim that our past-directed preferences are *always* practically inert. But he holds that when a future biased preference would influence the agent's choices, or would contradict an earlier preference on which she has already based a choice, future bias is rationally impermissible, and moreover is no longer psychologically typical.

on that preference, in conditions in which there is an action that can be taken (as would be the case if Parfit could bring it about that he has the 10-h past surgery).[10]

Assuming that preferences are evaluative mental states, those states will be sensitive to various factors: that is, various factors will influence how attractive a state of affairs is evaluated as being. And one such factor that might play a role is the relative degree to which events are causally manipulable (Latham et al., 2023).

Given our assumption that there is no actual backwards causation it follows that past events cannot be causally manipulated, and this might explain why people discount the value of past events relative to future events and thus explain why people are apparently future biased. Equally, far future events are often *less* causally manipulable than nearer ones. That is because, as Lewis (1979) notes, a cause typically has more effects than it does events that are its cause. While ordinary events often have myriad effects, massive overdetermination of some events by previous causes is rare. This is why our world contains causal forks. Given this, if we intervene on an event in the present time, it will have numerous effects that occur in the very near future as a result. And although there are multiple such effects, we have some chance of predicting what those effects will be, and hence some ability to bring about those effects by bringing about that cause. Over time, though, that single cause will have effects which will, in turn, serve as causes for further effects through time. That means that the present cause will have many effects indeed by the time we look at its effects in the far future. This makes it very difficult to predict *which* effects that cause will have in the far future, and makes it correspondingly difficult to causally intervene now, in such a way as to bring about some *particular* effect in the far future. Of course, this is not to say that we can never predict far future effects from present causes. We certainly can (think here of climate change). It is just to say that it is typically *harder* to predict the effects of some present cause in the far future compared to the nearer one. Thus, it is concomitantly harder to act now, so as to bring about some desired far future effect, than it is to bring about some desired near future effect. This might explain why we discount the value of far future events relative to nearer ones, and hence explain why we are apparently prospectively near biased.

However, since there is no difference in relative manipulability between past near events and past far events, it seems unlikely that any appeal to differences in causal manipulability could explain apparent retrospective near bias. So, in what follows we focus on apparent future bias and apparent prospective near bias.

According to the *manipulability hypothesis*, what (at least partially) explains apparent future bias and apparent prospective near bias is that our preferences are sensitive to the degree to which we can causally manipulate events.

We begin, in Sect. 2, by motivating the manipulability hypothesis and considering its normative implications, before outlining our more specific hypotheses and describing, in broad terms, how we test them. Then in Sect. 3, we outline our methodology and results in detail, before in Sect. 4 considering the implications of these results.

[10] However, there is recent work suggesting that future biased preferences do make a difference to action when combined with other preferences (see Dougherty, 2015; Greene & Sullivan, 2015, Braddon-Mitchell et al., 2023).

2 The manipulability hypothesis

There are several reasons to suppose that the manipulability hypothesis may find empirical support. We know that attention tends to be allocated towards stimuli that are relevant to attaining one's goals (Dijksterhuis & Aarts, 2010). Given that people can causally influence future events but not past ones, it will tend to be future events that are relevant to attaining their goals. In turn, people will tend to attend more to future than to past states of affairs. Likewise, if people can causally influence near future events more than far future events, these will tend to be more relevant to attaining their goals, and they will tend to attend to near future events more than further future events.

We also know that events over which we have control tend to elicit stronger emotional responses than those over which we lack control (Frijda, 1986; Frijda et al., 1989; Smith & Lazarus 1990). Given this, we would expect to find that people have a more intense affect with respect to future compared to past events, and to near future events compared to further future events.

In turn, the affective system guides decision-making in the pursuit of *appetitive* stimuli (i.e., something pleasant) and in the avoidance of *aversive* stimuli (i.e., something unpleasant). It also guides behaviour and decision-making by responding more strongly to goal-relevant stimuli, or to stimuli that require immediate action, such as immediate threats or immediately rewarding stimuli (Bradley et al., 2001; Bradley & Lang, 2007). So, we would expect the affective system to respond more strongly to future events than to past ones, and to near future events than to further future events.

Importantly, we know that attention and affect *jointly* guide behaviour and decision-making (Cacioppo et al., 1999; Mrkva et al., 2020). On the one hand, more attentional resources tend to be directed towards rewards or threats that are emotionally arousing (Lang et al., 1997). Thus, emotional intensity increases attention. Equally, stimuli to which we attend, tend to cause more emotional arousal than those to which we do not, and those stimuli tend to play more of a role in decision-making (Armel et al., 2008; Kahneman et al., 2006). So, if we respond more emotionally to future events than past ones, and our attention is partly determined by emotional arousal, then we should expect people to pay more attention to future events than past ones, which is in fact what we find (Caruso, 2010).

This attention asymmetry may have important consequences for our preferences, as there is evidence that people tend to value attended objects more, are more likely to value objects attended influencing value choices, and that increasing attention increases the likelihood of choice (Störmer & Alvarez, 2016; Pleskac et al., 2022; Ghaffari & Fiedler, 2018; Janiszewski et al., 2012). If people more highly value the option to which they attend, then attentional biases towards future compared to past may cause us to value an event more highly when it is future compared to when it is past, and thus prefer negative events to be past not future and positive ones to be future not past. Likewise, if people more highly attend to near future events than far future ones, then we would expect them to value an event more highly when it is in the near future compared to the far future, and thus prefer negative events to be in the far future and positive ones in the near future.

Thus, we'd expect that perceived differences in causal manipulability will produce both a difference in emotion and in attention between past/future and near future/far future events. In turn, given the time-asymmetrical connections between emotion and attention and choice preference, there is reason to think that this difference in emotion and attention will lead people to differently value events depending on where they are located. People will find past events less salient and less motivating towards action than future ones, and will find far future events less salient and less motivating towards action than near future ones. Further, they will find that the former evokes less intense affect than the latter, and because of this, we would expect people to value those events less because they *care* about them less. Valuing past events less than future events, and far future events less than near ones explains both apparent near bias and apparent future bias, since if someone values past events less than future ones they will prefer that positive events are located in the future (where they have a higher value) than in the past, and that negative events are located in the past rather than the future, where they have less negative value. *Mutatis mutandis* for near future versus far future events (Ramos et al., 2022).

Importantly, the manipulability hypothesis is consistent with the idea that people discount the value of events in direct proportion to the degree to which those events are causally manipulable so that events that are 100% manipulable are not discounted at all, while those that are 0% manipulable are 100% discounted, that is, are accorded no value. If that were right, then we would expect to find that people accord no value to past events whatsoever and less value to far future events than nearer ones. This, however, is not what we find when it comes to past events; in fact, people do accord such events some value (Latham et al., 2021a, Bickel et al., 2008; Yi et al., 2006; Greene et al., 2021a, b, 2022c).

The hypothesis is, however, also consistent with the idea that there is a discount function that discounts the value of events when they are less than 100% causally manipulable but does not discount them 100% even if they are not manipulable at all (in this case, the discount curve entirely flattens out when events are 0% manipulable but does not discount those events entirely). This kind of discount function is consistent with the evidence we already have observed regarding past discounting.

Moreover, we think that given empirical evidence, we would not expect people to entirely discount the value of past events even if they are not at all manipulable, given the role of emotion, salience, and attention and the ways these are connected to causal manipulability. Although past events are not causally manipulable, that does not mean that evolved agents would do well if those events had no salience, were emotionally neutral and garnered no attention. Agents make decisions about what to do based in part on information about the outcomes of their past actions. If the past pain, for instance, is salient because it is still emotionally arousing and attention-garnering, then an individual is more likely to take this into account in making future decisions. In general, creatures who do not direct any attentional or emotional resources towards past events would tend to be less motivated to avoid previously encountered bad outcomes, and less motivated towards previously encountered good outcomes. And so, we would expect there to be evolutionary pressure to direct some attentional and emotional resources towards past events. We would also expect that if the manipulability hypothesis were

correct, then people will discount the value of past events relative to future ones but will not discount their value entirely.

Even though many authors have endorsed the manipulability hypothesis, very little empirical work has been done in this regard. While many prior studies have tested the conditions under which people are apparently near or future biased, there is only one study we are aware of that has gone any way at all toward probing the manipulability hypothesis, which is a 2020 study by Latham et al. (2020). In that study they compared levels of apparent future bias regarding negative events, across two conditions. In the *choice* condition participants saw a vignette in which the protagonist can choose whether to undergo a single painful future shock or not. A predictor has time travelled from a past time to the present time to witness the choice. If they choose to undergo the additional future shock this will retrospectively cause it to be the case (via the predictor's actions) that in the past, they were administered fewer shocks than if they do not choose to undergo the additional shock. Thus, the protagonist can choose to bring it about that they suffer fewer shocks overall, by undergoing a single future shock. Or they can choose to bring it about that they had more shocks overall, but all of which are in the past, if they decline the future shock. In the *preference* condition, participants were asked whether, as the protagonist in a similar vignette, they prefer to be on a schedule in which they received more past shocks and no future one, or a schedule on which they receive one future shock but fewer past ones. Latham et al. (2020) hypothesised that people would show less apparent future bias in the choice condition compared to the preference condition and that is what they found. They took this to be evidence in favour of the idea that when the past is made causally relevant, apparent future bias is diminished.

There are, however, are several limitations to this study in terms of evaluating the manipulability hypothesis. First, that study compared a choice condition to a preference condition, where the former involved causal manipulability of the past, and the latter did not. We now know from more recent research that the capacity for choice *itself* decreases apparent future bias (Greene et al., 2022a, 2022b) even in the *absence* of any causal manipulability. As a result, their results may speak more to the nature of the role of choice compared to preference, than to the role of causal manipulability versus not. Second, although their study makes the past event causally manipulable, their vignettes are not ones in which some event, E, which can either occur in the past, or the future, is equally causally manipulable wherever it is. Rather, theirs were cases in which the protagonist could choose to perform a *future* action that would make the *past* some way or could choose *not* to perform that future action. This is notably different from standard cases used to elicit preferences about future bias. Think again about the case of Parfit's Operation. Parfit claims to prefer a much longer past operation to a shorter future one. But there may well be a difference between Parfit saying that if he could, he would causally bring it about that he had the longer past operation and will not have the shorter future one—which we take it would be a standard way to take an ordinary sort of future bias case and make the past causally manipulable—and Parfit saying that he would *not* bring it about that he has a very minor procedure now, to bring it about that he had a shorter operation in the past rather than a longer one, if he does not have the minor procedure now. Latham et al. (2020) tested vignettes of the latter kind; we aim to test vignettes of the former kind. Third, Latham et al.

(2020) focused only on negative events, so even if their results do speak to causal manipulability, they might not be fully general. Finally, Latham et al. (2020) only focused on apparent future bias and not on apparent near bias. Thus, in all these ways this single study of fails to provide good evidence regarding the status of the causal manipulability thesis. Our study aims to eliminate all these limitations to test whether in fact causal manipulability plays a role in explaining (some) apparent time biases.

Following Latham et al. (2020) we can distinguish two versions of the manipulability hypothesis. According to the *narrow adaptation* version we have an evolved general disposition to discount the value of further future events compared to nearer ones, and to discount the value of past events compared to future ones, and we evolved this general disposition *because* of the relative manipulability of these events. According to the *cognitive mediation* version of the hypothesis, our preferences are sensitive to our beliefs about the degree to which we can causally manipulate events, and we discount the value of events that we cannot manipulate or can manipulate less reliably.

Given the results found in Latham et al. (2020), there is some reason to suppose that insofar as the manipulability hypothesis is correct, it is the cognitively mediated version of that hypothesis that is correct. That is because Latham et al. found that people's beliefs about whether a particular past event was causally manipulable or not, made a difference to their preferences regarding that event, in a way that we would not expect if the narrow adaption view were correct. Henceforth then, it is the cognitively mediated version of the hypothesis that we shall be investigating.

The question then arises as to the normative implications of the manipulability hypothesis. We might think that if the manipulability hypothesis is vindicated, then insofar as it explains why we have these preferences it tends to confer on them the status of rational permissibility. We take it that a factor tends to confer rational permissibility on a preference if, in being sensitive to that factor, the preference is sensitive to something that is *normatively relevant.* For instance, if Freddie's preference for roo over watermelon is sensitive to the *flavour* of roo and watermelon, then the preference is sensitive to something normatively relevant; by contrast, if it is sensitive to whether the roo is to the *left*, or *right*, of the watermelon (from Freddie's perspective) then this is to be sensitive to something normatively irrelevant.

Next the question becomes whether being sensitive to the relative causal manipulability of an event is normatively relevant to a preference regarding that event's location. We cannot decisively answer that question here. One consideration that might be brought to bear in this regard, though, is to consider whether agents who are sensitive in this manner tend to do better than those who are not, in the sense that they tend to have a higher degree of well-being (whether this is defined in terms of having more of what is on some list of objective goods, or by having more of their preferences satisfied). We think that an argument can at least be mounted to the conclusion that this is the case. That is because, at least on the face of it, agents will typically do better in a whole host of ways if they attend to, and are motivated by, near future states of affairs more than past ones or far future ones. As Horwich (1987, p. 197) notes, "[A]n organism that wanted its future selfish desires to be satisfied would flourish relative to an organism that didn't care; however there is no particular advantage in wanting past desires to have been satisfied." There is no advantage to wanting past desires to be satisfied because in our world there is nothing we can do to bring about past

states of affairs. To put things as Ramos et al. do, "spending more time and effort on the future than the past may constitute an optimal pattern of resource allocation, and thus be highly functional." Similar thoughts have also been articulated by Hare (2013), Suddendorf and Corballis (2007), Maclaurin and Dyke (2002), and Suhler and Callender (2012).

Thus, it might be argued that because having preferences that are sensitive to the relative manipulability of events leaves one better off, in terms of well-being and desire satisfaction, then being sensitive to that factor is normatively relevant, and hence if our preferences are sensitive to this factor, then this tends to confer on them the status of being rationally permissible. Of course, there are several things to bear in mind here. First, even if this is all correct it does not entail that the resulting preference is, all things considered, rationally permissible, since it might also be the product of a bunch of *other* factors that are not normatively relevant. It would, however, suggest that such preferences are rationally permissible *insofar* as they are explained by this factor. Second, one might deny either that attending in this manner is beneficial to well-being, or that, if it is, then this tends to confer rational permissibility on the resulting preference. We take no stand on this matter in this paper; our point is just that determining whether the manipulability hypothesis is correct is important in terms of evaluating the normative status of these preferences.

Our aim is to empirically test the cognitive mediation version of the manipulability hypotheses.

So, there are two hypotheses we want to test.

Future (cognitively mediated) Manipulability Hypothesis (FMH): Our apparently future biased preferences are in part explained by our being sensitive to the degree to which we can causally manipulate events.
Near (cognitively mediated) Manipulability Hypothesis (NMH): Our apparently near biased preferences are in part explained by our being sensitive to the degree to which we can causally manipulate events.

If the FMH hypothesis is right, then we would expect that when past events are made as causally manipulable as future ones, people will discount the value of past events less, and therefore show lower levels of apparent future bias. So, we should find the following:

H1 When both past and future events are causally manipulable, we will find lower levels of apparent future bias and higher levels of apparent time neutrality.

If the NMH hypothesis is right, then if far future events are as manipulable as near ones, people should discount the value of those far future events less and therefore show lower levels of apparent near bias. Then we should find:

H2 When the far future is as causally manipulable as the near future, we will find lower levels of apparent near bias and higher levels of apparent time neutrality.

Our study tests H1 and H2.

To do this, we present participants with a vignette modified from Greene et al. (2021a) and Latham et al. (2023). In the vignette, participants are asked to *imagine*

that they receive either their most disliked meal (in the negative condition) or their favourite meal (in the positive condition), where the meal can either be received in the past or in the future (in the apparent future bias condition) or can be received in the near future or in the far future (in the apparent near bias condition). The vignette is contrived so that the meal is certain to be received, and the only question is *when* it will be received, and so that the person receiving it will find it *equally* tasty/unpleasant *whenever* it is received. Thus, the vignette is intended to control for both the intrinsic utility/disutility of the event, as well as its probability of occurring. This might seem puzzling in the case of past events, so it is worth pausing to say a bit more here.

Time biased preferences are preferences that people have when they are *unsure when* an event occurs and are asked to form preferences over the temporal location of the event. In the case of apparent near bias, this is straightforward. In the case of apparent future bias, however, it is not, since people often *know* whether an event occurred in the past or not, by remembering it. If an agent *remembers* that an event happened, then asking them whether they prefer it to be in the past, or the future, may seem puzzling. That is why apparent future bias scenarios, from the one that Parfit outlined in Parfit's Operation onwards, tend to be ones in which the agent is unsure whether an event already happened, or is yet to happen. Then people's preferences over these two possibilities are elicited. And the agent is in a position to attach a credence to the probability of the event happening at each temporal location. Typically, as in our vignettes, participants are told that an event *certainly* happens, and what is unknown is whether the event was in the past, or will be in the future, or whether it will be in the near future or far future. This is the sense in which the probability of the event is controlled in our study.

To see what it would be to fail to control for this, let's imagine some other cases. We could ask people to imagine a case in which the event is more likely to occur if it happens earlier, rather than later in the future. For instance, Freddie is more likely to get to eat the cake if he eats it earlier rather than waits till later, when someone else might have already eaten it. We can imagine similar cases when it comes to events in the past/future. Suppose we tell Freddie that he needs to undergo a painful operation. It can either be that he already underwent the operation, or that he will undergo the operation at some future time. While Freddie might construe that as a situation in which the operation is certain, and the only question is *when* it will be, he might also reason that if he already had the operation, then, of course, conditional on preferring a past operation he is preferring an operation that is certain to have occurred. On the other hand, he might reason that if he has not yet had the operation, then there is some chance that he may end up not needing it, and hence not having it. Perhaps for instance a cure will suddenly be discovered, or it will be discovered that he is not ill after all; perhaps he has just been misdiagnosed. Then he might suppose that if the operation was not in the past, then there is some chance that the operation will not occur at all. Thus, Freddie might compare the *certainly* of a past painful operation, with the mere *probability* of a future painful one. In such a case Freddie's subjective probability of undergoing the operation is not the same regardless of where it is located. Freddie clearly has a reason to prefer a less certain future operation over a certain past one, and that is why we aim to control for probability.

Moving on, in this vignette we make no mention of how causally manipulable the past is compared to the future or how manipulable the far future is compared to the near future. Rather, we allow participants to view the vignette as they would any case in which they are asked to form preferences over temporal locations. Notice that when people are *standardly* presented with vignettes to test for apparent future bias, they are not told that they *cannot* causally manipulate the past, and so in this regard, the vignettes are the same. We then ask participants whether they prefer that the meal is in the near as opposed to far future (in the apparent near bias condition), or in the future compared to the past (in the apparent future bias condition).

This gives us a record of people's preference in what we call *the standard conditions.*

A second set of participants is then presented with an amended version of each of these vignettes, in which we aim to make the relevant events equally causally manipulable. The vignettes are all in other respects the same: so, once again, they control for the probability of the event being the same whenever it is located, and for the event having the same utility/disutility whenever it is located.

In the apparent future bias condition, we aim to make the past event as causally manipulable as the future event. As we noted previously, we assume that in the actual world, agents are not able to causally manipulate past events. So, in the vignettes participants are being asked to *imagine* that they are a protagonist in a situation in which they *are* able to causally manipulate past events because they have a special 'backwards causation' device that allows them to do so. Thus, we are asking participants to imagine that things are not the way they likely take them to be, and asking what they do would under that supposition. However, since the idea of backwards causation is not especially foreign to people (think here of all of the myriad time travel films that people watch and understand) we take it that most people are able to *imagine* that they can causally manipulate past events and to decide what they would do in that eventuality. This is what our vignettes test.

To do this, we introduce an additional paragraph to the original vignette, which specifies that the individual in the vignette (who the participant is to imagine is themselves) has access to a backwards causation machine which has two buttons: yellow and blue. If the individual presses the yellow button, they bring it about that the event (the meal being dispensed) happened in the past, while if they press the blue button, they bring it about that the meal will be dispensed in the future. Thus, the individual has equal causal control over the dispensing of the meal. Participants are then asked which button they will press.

In the apparent near bias condition we amended the same vignette as was used in the original apparent near bias condition, but we added a paragraph that says that the individual in the vignette (who participants are to imagine is themselves) has discovered that the dispenser can be programmed so that it is *guaranteed* to bring about certain events at a particular future time, with 100% certainty. They are told that they, now, can either program the dispenser to deliver their food at a particular time in the near future, or at a particular time in the further future. Again, the aim is for participants to see that they have *equal control* of when the meal is dispensed, since the program is 100% guaranteed to deliver what it is programmed to deliver, and

participants have control over what is programmed. Participants are then asked what they program the machine to do.

This gives us a record of people's preference in what we call *the manipulability conditions*. These are conditions in which people have equal control over the event regardless of where in time it is located.

If the cognitive mediation version of the manipulability hypothesis is correct, we should find that people show more future bias and near bias in the standard conditions than in the manipulability conditions.

Below we present our experimental design in more detail, including the vignettes and the various comprehension and probe questions we asked participants.

3 Experimental design and results

3.1 Method

3.1.1 Participants

1635 people participated in the study. Participants were U.S. residents, recruited and tested online using Amazon Mechanical Turk (MTurk), and compensated $1 for approximately 5 minutes of their time. Given recent worries about the quality of data collected through MTurk, concerning both the quality of human responders and the presence of bots, we adopted several quality-control measures.[11] First, we used only those MTurk participants who had a HIT (task) approval rate of at least 95% and who have had their HITs (tasks) approved at least 1000 times. That means that all our participants had already successfully completed at least 1000 other tasks and received at least a 95% approval rating on these tasks, a standard that can be expected to eliminate most bots. Second, our study included both attentional checks and comprehension checks shown after each vignette. In total, 900 participants had to be excluded for failing to follow task instructions, or for failing to answer 3 comprehension questions and an attentional check question correctly. The remaining sample was composed of 753 participants (aged 20–79; 356 female, and 16 Trans/Non-binary; Mean age 42.15 (SD = 12.34)). Ethics approval for this study was obtained from the University of Sydney Human Research Ethics Committee. Informed consent was obtained from all participants prior to testing. The survey was conducted online using Qualtrics.

3.1.2 Materials and procedure

The experiment is a 2 × 2 × 2 between-subjects design with participants randomly assigned to one of eight conditions. The eight conditions reflect all possible combinations of (1) valence—positive or negative, (2) kind of preference—that is, whether they are being asked a preference for an event being in the past/future (testing for apparent future bias) or being in the near-future/far-future (testing for apparent near

[11] See Ahler et al. (2021) for a discussion of some of the problems associated with collecting data using MTurk and the prevalence thereof.

bias), and (3) degree of causal manipulability—standard causal manipulability of past or far-future vs. full causal manipulability of the past or far-future.

We developed a single base vignette that could be minimally modified for the eight conditions of the experiment. The vignette is a modified version of the vignette used in Greene et al. (2021a) and Latham, Miller and Norton (forthcoming). All participants read a vignette describing a hedonic event (receiving a meal) which can either be a favourite or most disliked meal. The modifications include valence, the kind of scenario, and the degree of causal manipulability. The difference between the positively and negatively valenced conditions is indicated in square brackets.

The attention question for all vignettes was: In this vignette, you were asked to imagine that you were...

(a) A food dispenser
(b) A spaceship.
(c) A dog.
(d) An astronaut.

Experiment vignettes

Future Bias, Base Case:

Imagine you are an astronaut on a 10-year voyage between planets. You are half-way through your journey. Fortunately for you, space travel is extremely safe, and you are completely certain that you will survive your voyage.

The ship's food dispenser normally produces bland meals containing only essential nutrients. However, the food dispenser is programmed to dispense, on one day of the voyage, the meal which is your [favourite]/[most disliked] meal on that day. Whenever the meal is dispensed, it will taste good to you, because the food dispenser is able to detect what you like, and do not like, at a particular time, and it will create the meal accordingly. So, for instance, if your tastes change over the course of the journey, the machine will dispense your [favourite]/[most disliked] meal at the time at which it is dispensed.

The ship's food dispenser is extremely reliable: in over 10,000 space journeys that have taken place over 1 million days of travel, a machine has never broken down or misjudged the culinary preference of an astronaut. So you are completely sure that it will dispense food every day, and that it will dispense your [favourite]/[most disliked] meal on one day.

As it happens, if the food dispenser has already dispensed your [favourite]/[most disliked] meal, then it dispensed your [favourite]/[most disliked] meal 3 times on one day 12 months ago. If it has not already dispensed your [favourite]/[most disliked] meal, then it will dispense it on one day in 12 months' time and it will dispense your [favourite]/[most disliked] meal only 1 time on that day.

One morning, you awake from a dream concerning your [favourite]/[most disliked] meal and for a moment you cannot remember whether you have received it yet.

Comprehension question 1: In this vignette, you were asked to imagine that during the 10-year voyage, the ship's food dispenser produced bland meals…

(a) One day a year.
(b) Every day.
(c) Every day except for one.
(d) One day a week.

Comprehension question 2: In this vignette, you were asked to imagine that if the food dispenser dispensed your [favourite]/[most disliked] meal 12 months ago, then…

(a) You received that meal 3 times during a year.
(b) You received that meal 3 times during one day.
(c) You received that meal 1 time during one day.
(d) You received that meal twice during a week.

Comprehension question 3: In this vignette, you were asked to imagine that if the food dispenser will dispense your [favourite]/[most disliked] meal in 12 months' time, then…

(a) You will receive that meal twice during a week.
(b) You will receive that meal 3 times during one day.
(c) You will receive that meal 3 times during the next year.
(d) You will receive that meal 1 time during one day.

Probe question: Please indicate your PREFERENCE using one of the following statements:

(a) I would prefer that my [favourite/most disliked] meal will be dispensed one time on one day in 12 months' time, and was not dispensed three times on one day 12 months ago.
(b) I would prefer that my [favourite/most disliked] meal will be dispensed three times on one day in 12 months ago, and was not dispensed one time on one day in 12 months' time.
(c) you have no preference between these options.

Future Bias, Causal Accessibility:

Imagine you are an astronaut on a 10-year voyage between planets. You are half-way through your journey. Fortunately for you, space travel is extremely safe, and you are completely certain that you will survive your voyage.

The ship's food dispenser normally produces bland meals containing only essential nutrients. However, the food dispenser is programmed to dispense, on one day of the voyage, the meal which is your [favourite/most disliked] meal on that day. Whenever the meal is dispensed, it will taste good to you, because the food dispenser is able to detect what you like, and do not like, at a particular time, and it will create the meal accordingly. So, for instance, if your tastes change over the course of the journey, the machine will dispense your [favourite/most disliked] meal at the time at which it is dispensed.

The ship's food dispenser is extremely reliable: in over 10,000 space journeys that have taken place over 1 million days of travel, a machine has never broken down or misjudged the culinary preference of an astronaut. So you are completely sure that it will dispense food every day, and that it will dispense your [favourite/most disliked] meal on one day.

As it happens, if the food dispenser has already dispensed your [favourite/most disliked] meal, then it dispensed your [favourite/most disliked] meal 3 times on one day 12 months ago. If it has not already dispensed your [favourite/most disliked] meal, then it will dispense it on one day in 12 months' time and it will dispense your [favourite/most disliked] meal only 1 time on that day.

On the spaceship you have access to a backwards causation machine. This machine has two buttons, a yellow button and a blue button. Pressing these buttons causes certain events to have happened in the past and in the future. If you press the yellow button, you bring it about that the food dispenser dispensed your [favourite/most disliked] meal on one day 12 months ago. So, you bring it about that your [favourite/most disliked] meal was dispensed 3 times on one day in the past. If you press the blue button you bring it about that the food dispenser did not already dispense your [favourite/most disliked] meal and will dispense it on one day in 12 months' time. So you bring it about that the dispenser will dispense your [favourite/most disliked] meal 1 time on one day in the future. If you press the blue button you will receive your [favourite/most disliked] meal only 1 time in the future. If you do not press either button, then the machine will choose at random which of these outcomes to bring about.

One morning, you awake from a dream concerning your [favourite/most disliked] meal and for a moment you cannot remember whether you have received it yet.

Comprehension question 1: In this vignette you were asked to imagine that during the 10-year voyage, the ship's food dispenser produced bland meals…

(a) One day a year.
(b) Every day.
(c) Every day except for one.
(d) One day a week.

Comprehension question 2: In this vignette you were asked to imagine that if you press the yellow button, then…

(a) You cause it to be the case that you will receive your [favourite/most disliked] meal 3 times during one day in 12 months' time.
(b) You cause it to be the case that you received your [favourite/most disliked] meal once during one day 12 months ago.
(c) You cause it to be the case that you will receive your [favourite/most disliked] meal once during one day in 12 months' time.
(d) You cause it to be the case that you received your [favourite/most disliked] meal 3 times during one day 12 months ago.

Comprehension question 3: In this vignette you were asked to imagine that if you press the blue button, then…

(a) You cause it to be the case that you will receive your [favourite/most disliked] meal 3 times during one day in 12 months' time.
(b) You cause it to be the case that you received your [favourite/most disliked] meal 3 times during one day 12 months ago.
(c) You cause it to be the case that you will receive your [favourite/most disliked] meal once during one day in 12 months' time.
(d) You cause it to be the case that you received your [favourite/most disliked] meal once during one day 12 months ago.

Probe question: Please indicate your CHOICE using one of the following statements.

(a) I would press the blue button and bring it about that my [favourite/most disliked] meal will be dispensed once, on one day in 12 months' time.
(b) I would press the yellow button and bring it about that my [favourite/most disliked] meal was dispensed three times on one day 12 months ago.
(c) I would choose to press neither button because I have no preference between the two options.

Near Bias, Base Case:

Imagine you are an astronaut on a 10-year voyage between planets. Fortunately for Imagine you are an astronaut on a 10-year voyage between planets. You are half-way through your journey. Fortunately for you, space travel is extremely safe, and you are completely certain that you will survive your voyage.

The ship's food dispenser normally produces bland meals containing only essential nutrients. However, the food dispenser is programmed to dispense, on one day of the voyage, the meal which is your [favourite/most disliked] meal on that day. Whenever the meal is dispensed, it will taste good to you, because the food dispenser is able to detect what you like, and do not like, at a particular time, and it will create the meal accordingly. So, for instance, if your tastes change over the course of the journey, the machine will dispense your [favourite/most disliked] meal at the time at which it is dispensed.

The ship's food dispenser is extremely reliable: in over 10,000 space journeys that have taken place over 1 million days of travel, a machine has never broken down or misjudged the culinary preference of an astronaut. So you are completely sure that it will dispense food every day, and that it will dispense your [favourite/most disliked] meal on one day.

As it happens, if the food dispenser will dispense your [favourite/most disliked] meal on one day in one week's time then it will dispense your [favourite/most disliked] meal only 1 time on that day. If it will dispense your [favourite/most disliked] meal on one day in 12 months' time, then it will dispense your [favourite/most disliked] meal 3 times on that day.

One morning, you awake from a dream concerning your [favourite/most disliked] meal and for a moment you cannot remember whether you have received it yet.

Comprehension question 1: In this vignette you were asked to imagine that during the 10-year voyage, the ship's food dispenser produced bland meals…

(1) Every day.
(2) One day a week.
(3) One day a year.
(4) Every day except for one.

Comprehension question 2: In this vignette you were asked to imagine that if the machine dispenses your [favourite/most disliked] meal on one day in one week's time, then…

(a) You will receive your [favourite/most disliked] meal twice during that day.
(b) You will not receive your [favourite/most disliked] meal at all.
(c) You will receive your [favourite/most disliked] meal three times during that day.
(d) You will receive your [favourite/most disliked] meal once during that day.

Comprehension question 3:: In this vignette, you were asked to imagine that if the machine dispenses your [favourite/most disliked] meal on one day in 12 months' time, then…

(a) You will receive your [favourite/most disliked] meal three times during that day.
(b) You will receive your [favourite/most disliked] meal twice during that day.
(c) You will receive your [favourite/most disliked] meal once during that day.
(d) You will not receive your [favourite/most disliked] meal at all.

Probe question: Please indicate your PREFERENCE using one of the following statements.

(a) I would prefer that my [favourite/most disliked] meal will be dispensed once on one day in one week's time, and will not be dispensed three times on one day in 12 months' time.
(b) I have no preference between these options.
(c) I would prefer that my [favourite/most disliked] meal will be dispensed three times on one day in 12 months' time, and will not be dispensed once on one day in one week's time.

Near Bias, Causal Accessibility:

Imagine you are an astronaut on a 10-year voyage between planets. You are half-way through your journey. Fortunately for you, space travel is extremely safe, and you are completely certain that you will survive your voyage.

The ship's food dispenser normally produces bland meals containing only essential nutrients. However, the food dispenser is programmed to dispense, on one day of the voyage, the meal which is your favourite meal on that day. Whenever the meal is dispensed, it will taste good to you, because the food dispenser is able to detect what you like, and do not like, at a particular time, and it will create the meal accordingly. So, for instance, if your tastes change over the course of the journey, the machine will dispense your favourite meal at the time at which it is dispensed.

The ship's food dispenser is extremely reliable: in over 10,000 space journeys that have taken place over 1 million days of travel, a machine has never broken down or misjudged the culinary preference of an astronaut. So you are completely sure that it will dispense food every day, and that it will dispense your favourite meal on one day.

As it happens, if the food dispenser will dispense your favourite meal on one day in one week's time then it will dispense your favourite meal only 1 time on that day. If it will dispense your favourite meal on one day in 12 months' time, then it will dispense your favourite meal 3 times on that day.

You have just found out today that the food dispenser can be programmed so that it is guaranteed to bring about certain events at a particular future time with 100% certainty. You can either program the food dispenser to bring it about that your favourite meal will be dispensed 1 time on a day in one week's time, or to bring about that your favourite meal will be dispensed 3 times on a day in 12 months' time.

Comprehension question 1: In this vignette, you were asked to imagine that during the 10-year voyage, the ship's food dispenser produced bland meals…

(a) Every day.
(b) One day a week.
(c) One day a year.
(d) Every day except for one.

Comprehension question 2: In this vignette, you were asked to imagine that if you program the machine to dispense your favourite meal with 100% certainty in one week's time, then…

(a) You will receive your favourite meal three times during that day.
(b) You will not receive your favourite meal at all.

(c) You will receive your favourite meal once during that day.
(d) You will receive your favourite meal twice during that day.

Comprehension question 3: In this vignette, you were asked to imagine that if you program the machine to dispense your favourite meal with 100% certainty in 12 months' time, then…

(a) You will receive your favourite meal twice during that day.
(b) You will not receive your favourite meal at all.
(c) You will receive your favourite meal once during that day.
(d) You will receive your favourite meal three times during that day.

Probe question: Please indicate your CHOICE using one of the following statements.

(a) I do not care which setting the food dispenser is programmed to.
(b) I would set the food dispenser to dispense my [favourite/most disliked] meal three times on one day in 12 months' time.
(c) I would set the food dispenser to dispense my [favourite/most disliked] meal once on one day in one week's time.

4 Results and analysis

Before describing our analyses, we will begin by summarising our findings. Neither of our hypotheses were vindicated. H1 predicted that there would be lower levels of apparent future bias in causal manipulability conditions compared to standard conditions. While we observed an association between participants' apparent future biased preferences and causal manipulability, this was due to participants being relatively more likely to be apparently past biased in standard conditions than in causal manipulability conditions. H2 predicted that there would be lower levels of apparent near bias in causal manipulability conditions compared to standard conditions. We found *no* evidence of an association between participants' apparent near biased preferences and causal manipulability.

Table 1 below summarises the descriptive data of participants' reported future bias in the experiment. The 'FB' column represents the number of participants who report positive and negative apparent future biased preferences. A positive future biased

Table 1 Descriptive data from all conditions in the experiment of participants' apparent future biased preferences

Condition	FB	PB	TN
Positive event			
Standard ($n = 75$)	36 (48.0%)	24 (32.0%)	15 (20.0%)
Causal ($n = 60$)	30 (50.0%)	12 (20.0%)	18 (30.0%)
Negative event			
Standard ($n = 47$)	18 (38.3%)	25 (53.2%)	4 (8.5%)
Causal ($n = 33$)	22 (66.7%)	8 (24.2%)	3 (9.1%)

preference means that you prefer positive events to be in the future, whereas a negative future biased preference means that you prefer negative events to be in the past. The 'PB' column represents the number of participants who report apparent positive and negative past biased preferences. A positive past biased preference means that you prefer positive events to be in the past, whereas a negative past biased preference means that you prefer negative events to be in the future. The 'TN' column represents the number of participants who report apparent positive and negative time neutral preferences. A positive or negative time neutral preference means that you have no preference whether positive or negative events be in the future or the past. To test whether there was any association between participants' apparent future biased preferences and conditions we performed a chi-squared test of homogeneity. The test showed that there was a significant association between condition and participants reported apparent future biased preference, $\chi^2(6, N = 215) = 21.537, p = .001$).

Next, to check whether there was any association between participants' apparent future biased preferences and valence we ran a chi-square test of homogeneity. This test showed that there was a significant association between participant's apparent future biased preferences and valence, $\chi^2(2, N = 215) = 9.992, p = .007$). Post-hoc comparisons with a Bonferroni correction showed that participants were more likely to report being apparently time neutral in positive conditions than in negative conditions, $p = .004$.

Finally, to test whether there was any association between participants' apparent future biased preferences and causal manipulability we ran another chi-square test of homogeneity. This test showed that there was a significant association between participants' apparent future biased preferences and causal manipulability, $\chi^2(2, N = 215) = 8.570, p = .014$). Post-hoc comparisons with a Bonferroni correction showed that participants were more likely to report being apparently past biased in standard conditions than in causal accessibility conditions, $p = .004$.

Table 2 below summarises the descriptive data of participants' reported NB preferences in the experiment. The 'NB' column represents the number of participants who report positive and negative apparent near biased preferences. A positive near biased preference means that you prefer positive events to be temporally near, whereas a negative near biased preference means that you prefer negative events to be temporally far away. The 'FrB' column represents the number of participants who report positive and negative apparent far biased preferences. A positive far biased preference means

Table 2 Descriptive data from all conditions in the experiment of participants' apparent future biased preferences

Condition	FB	PB	TN
Positive event			
Standard ($n = 75$)	36 (48.0%)	24 (32.0%)	15 (20.0%)
Causal ($n = 60$)	30 (50.0%)	12 (20.0%)	18 (30.0%)
Negative event			
Standard ($n = 47$)	18 (38.3%)	25 (53.2%)	4 (8.5%)
Causal ($n = 33$)	22 (66.7%)	8 (24.2%)	3 (9.1%)

that you prefer positive events to be temporally far away, whereas a negative far biased preference means that you prefer negative events to be temporally near. Finally, the 'TN' column represents the number of participants who report positive and negative apparent time neutral preference. A positive or negative time neutral preference means that you have no preference whether positive or negative events are temporally near or far. To test whether there was any association between participants' apparent near biased preferences and condition we performed a chi-square test of homogeneity. The test showed that there was a significant association between condition and participants reported apparent near biased preferences, $\chi^2(6, N = 273) = 14.800$, $p = .022$. Next, to check whether there was any association between participants' apparent near biased preferences and valence we ran a chi-square test of homogeneity. This test showed that there was a significant association between participants' apparent near biased preferences and valence, $\chi^2(2, N = 273) = 11.981$, $p = .003$). Post-hoc comparisons with a Bonferroni correction showed that participants were more likely to report being apparently near biased in positive conditions than in negative conditions, $p < .001$, and report being apparently far biased in negative conditions than in positive conditions, $p = .002$.

Finally, to test whether there was any association between participants' apparent near biased preferences and causal manipulability we ran another chi-square test of homogeneity. The test found *no* evidence of an association between participants' apparent near biased preferences and causal manipulability, $\chi^2(2, N = 273) = 1.902$, $p = 0.386$).

5 Discussion

There are several notable features of our results. First, although we made no predictions in this regard, our findings replicated earlier findings regarding the role of valence. Greene et al. (2022c, 2022d) found that future bias is stronger when it comes to negative events than positive ones, a finding replicated in Latham et al. (2023). Likewise, studies on future temporal discounting have shown that people discount future positives more than future negatives (Mischel et al., 1969; Read, 2004; Thaler, 1981; Latham et al., 2024 Lh. Our results also replicated these findings.

However, neither of our hypotheses was vindicated. We found no evidence in favour of the cognitive mediation version of the manipulability hypotheses. Interestingly, however, we did find higher levels of apparent *past bias* in the standard condition compared to the manipulability condition. Past bias is the inverse of future bias: it is a preference to have negative events in the future, not the past, and positive ones in the past not the future, all else being equal. So, while when we made the past as causally manipulable as the future this did not *decrease* apparent future bias as predicted, we did find lower levels of apparent *past* bias in this condition compared to the standard condition. This suggests there may be *some* connection between manipulability and apparent future bias.

To see what that connection might be, consider the explanation for the presence of apparent past bias found in several experiments by Greene et al. (2021a, 2021b). They suggest that when people perceive it to be within their power to mitigate the badness

of some event if it is future, they are more inclined to prefer to locate that event in the future rather than the past. They call this the *mitigation hypothesis*. So, people will sometimes show apparent past bias because they are tacitly comparing a less bad future event with a worse past event, and preferring the former to the latter. If they are right about this, then we might expect to find that when the past is as manipulable as the future, people will show *decreased* levels of past bias. For it will no longer be true that negative events can only be mitigated if they are in the future, and hence there will be no reason to prefer those events to be future rather than past. This is just what we found.

One might worry, however, that if this account is right, we should also expect that manipulability will have an effect on apparent near/far bias. Since we find no such effect, this might be taken to undermine this explanation in terms of mitigation.

In fact, though, we think it is far from obvious *what* effect mitigation reasoning can be expected to have on apparent near and far bias (where apparent far bias is the inverse of apparent near bias). On the one hand, it might be that mitigation reasoning tends to promote apparent far bias over apparent near bias since temporally nearer events are typically more manipulable. As such, if we want to manipulate the badness of a bad event, we do better to have it located nearer rather than further, and hence to be apparently far biased. If this is right, then we should expect to find more apparent far bias in the manipulation condition (something we did not find). On the other hand, one might also think that the more time one has to plan for an event, the more one can mitigate its badness, and that this suggests that mitigation reasoning will lead us to be more apparently near biased: for we will prefer that negative events are in the farther rather than nearer future.

Given this, we think that the effect of manipulation on apparent near bias, insofar as it is moderated by manipulation reasoning, could be expected to be much less than the effect on apparent future bias. First, the case of apparent near bias mitigation reasoning is plausibly going to be sensitive to multiple factors (manipulation and time) that pull in opposite directions. Second, while past events are unmanipulable, both near and future events are usually manipulable to some degree, even if not to the same degree. Thus, we'd expect that making past events as manipulable as future ones will have a bigger impact of mitigation reasoning than will making far future events as manipulable as nearer ones. Thus, it could be that our results are in part explained by the kind of mitigation reasoning that Greene et al. (2022a, 2022b) propose. Insofar as such preferences are informed by this kind of reasoning, this would tend to suggest that they are rationally permissible. After all, it seems rationally permissible to prefer that, overall, one has less disutility and more utility. Mitigation reasoning aims to decrease disutility and increase utility by locating negative events in the future where their badness can be mitigated.

Still, our results did not replicate the findings of Latham et al. (2020) who found that causal manipulability decreases apparent future bias. This is puzzling.

There are several differences between the vignettes used in Latham et al. (2020) and those used in our study, which might explain these different results. As we noted earlier in the paper, in that study the vignettes in which there is causal manipulation of the past are also ones in which there is choice, while those without causal manipulation are ones in which people are merely asked their preference. Thus, the study does

not isolate causal manipulability from choice/preference. Since we know from prior studies that framing vignettes in terms of exercising a choice rather than forming a preference makes a difference to the degree to which people display future bias, the difference found in this study may reflect the difference between choice and preference rather than between the ability to causally manipulate versus not. Second, as we also noted earlier, the setup of the vignettes in our study is much closer to a standard future bias vignette, while the setup in the Latham et al. (2020) study is notably different. This too might explain the difference in our results.

Third, it might be that the differences we see are due to the role of memory. The act of recollecting an event itself can be a pleasurable or unpleasant experience (Elster & Loewenstein, 1992). That a memory is very unpleasant is a reason to prefer that the event (that the memory is about) either be in the future (particularly the far future) where it cannot, yet, be remembered, or a reason to prefer that the past event be less bad (and hence the memory be less bad). Indeed, the idea that memory plays a role in explaining apparent future bias has been suggested by several authors. If the pleasure or pain of recalling a memory is greater than that associated with some future event, then people might prefer more pleasure in the past to less in the future, and to prefer less pain in the past, to more in the future, since that might still result in more overall pleasure and less overall pain in the future, taking into account the pleasure/pain of remembering. One recent study by Lee et al. (2022) tends to support this idea. They found that a significant number of adults preferred 10 painful treatments in the past compared to one in the future when they would have no memory of the treatments but preferred one painful treatment in the future to 10 in the past when they would regain their memory. As such it seems that for some adults, the disutility of remembering a substantially painful past event outweighs the disutility of living through a less painful future event.

We assume that a memory of a series of very painful shocks would be very unpleasant. It may be, then, that people were motivated to reduce the number of past shocks experienced to reduce the unpleasantness of their memory of those experiences. By contrast, people are not likely to suppose that they will carry a painful and unpleasant memory of having eaten bad food, and so they may be less motivated to try to reduce the badness of the past event by instead choosing a less bad future event. That is, in the case of the meals people may have judged that the relief that the three bad meals are over and not being anticipated, not only outweighed the increased badness of the meals (three of them versus one) but also outweighed the fact that there would be a somewhat unpleasant memory of having eaten the three meals. In contrast, people may have judged that the extreme unpleasantness of a memory of being shocked, combined with the additional past shocks being extremely bad, outweighed the fact that to diminish those shocks they would need to choose one additional future shock (a shock that would be administered very soon, and hence would generate very little negative waiting). Thus, this difference in memory effects might also explain the different results across the two studies.

Finally, there is a difference in the stakes involved. In the Latham et al. (2020) study participants could bring it about that they were given significantly fewer painful electrical shocks in the (near) past, by electing to choose to have one shock in the (near) future. We take it that disliked meals (even three of them) are much less unpleasant

than a series of painful electrical shocks. It is plausible that people are more motivated to avoid past painful shocks in a way that they are not motivated to avoid past awful meals. Though having said that, people may also be more motivated to avoid a future painful shock than future disliked meals. Still, as we noted earlier in the paper, part of the proposed explanation of apparent near and future bias appeals to the role of the emotional intensity and salience of events and to the role of attention. It may be that events in the study by Latham et al. (2020) produced more emotional intensity than did our study and that this resulted in relative manipulability playing a role in influencing preferences.

One avenue of further investigation, then, would be to run a follow-up study that attempts to determine which of these explanations for this difference in results is the correct one.

As it stands, however, we did not find evidence in favour of the cognitive mediation version of the manipulability hypotheses. We think there are two possible conclusions to draw from this. First, it could be that although the cognitive mediation version of the hypothesis is not correct, the narrow adaptation version of the hypothesis is correct. That is, it could be that relative causal manipulability does play a role in explaining both apparent near and future bias, but that this effect is not cognitively mediated. Instead, it might be that people have a rough heuristic that involves discounting the value of past events compared to future ones and discounting the value of further future events compared to nearer ones and that these tendencies evolved because of the relative manipulability of past versus future, and near versus far, events. Yet this tendency might not be cognitively mediated, and hence people's preferences may not be sensitive to the relative manipulability of events. This would explain why people continue to be apparently future biased even when the past is made manipulable and continue to be apparently near biased even when the near and far future are equally manipulable. We think that further work in investigating this version of the manipulability hypothesis could profitably be undertaken.

A second possibility is that the causal manipulability hypotheses are false and that in fact, people's preferences are sensitive to some other factor which at least sometimes coincides with causal manipulability. One possibility is that what is *really* explaining people's apparent future and apparent prospectively near biased preferences is not the relative degree of manipulability per se of those events, but rather, people's judgements about how *probable* those events are. Subjective probability is often connected to manipulability. If we can causally bring about some event, then if we intend to do so, this will typically raise our credence that the event will occur. Thus, it will often be the case that not only is the near future more causally manipulable, but, in virtue of that, we have higher credence that an event will occur if it would be in the near future compared to if it would be in the farther future. Then, the thought is, if people's preferences are sensitive to how probable they take an event to be, then it may sometimes *appear* as though they are sensitive to how causally manipulable those events are, since the latter often accompanies the former.

To a limited extent our results are consistent with this idea. That is because, as we noted earlier, our vignettes control for subjective probability. They are cases in which participants are sure that an event will occur but are merely unsure when it will occur. If people are only sensitive to probability and not to causal manipulability, then they

will not be sensitive to manipulability *in the absence of probability differences.* Since we control for probability, our findings are consistent with the idea that it is probability, not manipulability, which people's preferences are sensitive to.

Having said that, it seems to us unlikely that if our preferences were sensitive to probability, that this would explain our displaying the apparently future biased preferences we do. There are only four ways our probability judgements could go. On one of them, we take the event in question to be certain regardless of when it is. In such a case if our preferences are sensitive to probability, them being so will not explain our having apparently future biased preferences. Similarly, we might judge that the event is certain not to occur, regardless of when it would be (or not be, in this case). Again, if this were so, then it would not explain our having apparently future biased preferences. So, the two interesting cases are ones in which we are either forming a preference over a more probable past event compared to a less probable future one or forming a preference over a more probable future event compared to a less probable past one. For instance, Freddie might take it that there is a 70% chance that he underwent a painful past operation, but that there is only a 60% chance that he will undergo a painful future operation. By contrast, he might take it that there is an 80% chance that he underwent a painful past operation, but a 90% chance that he will undergo a future painful operation. If Freddie's probability judgements are typically more like the latter than the former, then that can explain why he is apparently future biased, since obviously a smaller chance of having a painful operation is better than a larger one. And of course, if the event in question is a positive one rather than a negative one, then it needs to be that Freddie takes that event to be more probable if it is future compared to past so that he is comparing a more probable future positive event over a less probable past one.

The problem is that we see little reason to suppose that Freddie's probability judgments will typically, or at least more often, be such that he accords a *lower* subjective probability to some past event than to the same event in the future. Now of course, in *general* that will obviously not be true, since Freddie knows a good deal more about what did happen in the past, compared to what will happen in the future, since he has a lot of records of past events (including his own memories). So, in general we would expect him to be more certain about what did happen in at least his local past, compared to what will happen in his local future. But, recall, we are only focussing on events that Freddie is, in the present, unsure *when* (or indeed whether) they did/will occur. Even given this, though, it does not seem likely that in cases in which Freddie does not know whether an event occurred or is yet to occur, that he will *typically* think that such events are less likely to have occurred if they are in the past, compared to the future. And if that is right, then even if his preferences are sensitive to probability, this will not explain why he is apparently future biased. So, while perhaps apparent near bias is (partially) explained by people being sensitive to probabilities rather than causal manipulability, we think it unlikely that this explains apparent future bias.

Finally, it is worth reflecting on some limitations of our studies. One might worry that the vignettes were too cognitively demanding. To address this concern, we included multiple attention checks and comprehension check questions. This served two purposes. First, it weeded out bots, and people selecting answers at random and without thought to quickly receive payment (Ahler et al. 2021). This is something that

needs to be especially guarded against when running online studies and can result in large numbers of participant exclusions. In the case of these studies, we excluded 900 participants, which was almost half. Although this is a substantial rate of exclusion it falls below that in many other similar recent online studies (see for instance Latham et al., 2021a, 2021b) and is standard for MTurk studies in which weeding out people who did not pay attention or who are not people at all, is required.

Even with comprehension checks in play, however, one might worry that the vignettes are in some way 'unordinary' because they have a science fiction aspect, or because they include foreign concepts such as backwards causation that people may not have properly understood. The vignettes do have a science fiction aspect, insofar as they ask people to imagine that they are an astronaut on a spaceship. However, while not many people are in fact astronauts, we do not think that this requires a large imaginative leap; people frequently engage with fiction of this kind, and do not seem to have any difficulties with it. And this is the limit of what we ask people to imagine; we do not, for instance ask people to imagine that they are solving complex physics problems, or that they understand how a warp drive works. Moreover, the vignettes we used are minimally modified from vignettes used in previous studies, which found similar results when it comes to time bias (see Latham et al., 2023; Greene et al., 2021a, 2021b, 2022a, 2022b). Furthermore, other studies of time biases have used vignettes that do not have this science fiction aspect (see Latham et al., 2023; Caruso et al., forthcoming; Baron et al. 2023; Lee et al., 2022; Lee et al., 2020) and have found similar results regarding future bias.[12] Given that our results here, are consistent with previous results regarding time biases, we think there is little reason to be concerned that people did not understand, or could not simulate, being the protagonist in the vignette.

Having said that the manipulability vignettes do have an additional element, namely, in some of them, backwards causation. This does raise the possibility that participants did not really understand what they were being asked to imagine in those vignettes, perhaps because the idea that one can act in the present, to bring something about in the past, is foreign. We do think this is a potential concern. There are only two ways to make past events as causally manipulable as future ones: one is to make past events causally manipulable, and the other is to make future events not causally manipulable. We tend to think that since people are in fact familiar with backwards causation from an array of fiction, especially time travel fiction, that imagining that some past event can be influenced by the present, is no odder than imagining some future event that *cannot* be causally influenced by the present. In general, we would be surprised if people were unable to imagine backwards causation. Indeed, the previous work of Latham et al. (2020) who used quite complicated vignettes with time travelling predictors, suggests that people can imagine this, since in that study the authors found that the presence of those time travelling predictors did make a difference to people's preferences. We do, however, think that follow up work that attempted to probe causal manipulability in other ways, using other vignettes, would be worthwhile.

[12] The exact percentage of people who show future bias is sensitive the nature of the event in question, and the exact framing of the vignette, so there is certainly some variation across these studies (some of which also include children). However, what they show, taken as a corpus, is that people do display future bias across a wide range of cases, including those that do not have any science fiction element.

As things stand, however, we can say that we found no evidence of the cognitively mediated version of the manipulation hypothesis.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Declarations

Conflicts of interest The authors declare that they have no affiliations with any organization of any financial interest that would impact the content of the manuscript, and no other conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Ahler, D., Roush, C., & Sood, G. (2021). The micro-task market for lemons: Data quality on Amazon's Mechanical Turk. *Political Science Research and Methods, 8*, 1–20.

Ainslie, G. (2005). Précis of Breakdown of Will. *Behavioral and Brain Sciences, 28*(5), 635–650. https://doi.org/10.1017/S0140525X05000117

Baron, S., Everett, B. C., Lathan, A. J., Miller, A. J. H., Lathan, H., & Oh, J. V. T. (2023). Moving ego versus moving time: investigating the shared source of future-bias and near-bias. *Synthese*. https://doi.org/10.1007/s11229-023-04286-0

Bickel, W. K., Yi, R., Kowell, B. P., & Gatchalian, K. M. (2008). Cigarette smokers discount past and future rewards symmetrically and more than controls: is discounting a measure of impulsivity. *Drug and Alcohol Dependence, 96*, 256–262.

Braddon-Mitchell, D., Latham, A. J., & Miller, K. (2023). Can we turn people into pain pumps? On the rationality of future bias and strong risk aversion abstract.*Journal of Moral Philosophy, 1*, 1–23. https://doi.org/10.1163/17455243-20234084

Bradley, M. M., Codispoti, M., Cuthbert, B. N., & Lang, P. J. (2001). Emotion and motivation I: Defensive and appetitive reactions in picture processing. *Emotion (Washington, D.C.), 1(3)*, 276–298.

Bradley, M. M., & Lang, P. J. (2007). Emotion and motivation. In *Handbook of psychophysiology*, 3rd ed (pp. 581–607). Cambridge University Press. https://doi.org/10.1017/CBO9780511546396.025

Brink, D. O. (2010). Prospects for temporal neutrality. 10–011.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76*(5), 839–855. https://doi.org/10.1037/0022-3514.76.5.839

Caruso, E. M. (2010). When the future feels worse than the past: A temporal inconsistency in moral judgment. *Psychological Science, 139*(4), 610–624. https://doi.org/10.1037/a0020757

Caruso, E. M., Gilbert, D. T., & Wilson, T. D. (2008). A wrinkle in time: Asymmetric valuation of past and future events. *Psychological Science, 19*(8), 796–801.

Caruso, E., Latham, A. J., and Miller, K. (forthcoming). "Is future bias a manifestation of the temporal value asymmetry?" *Philosophical Psychology.*

Critchfield, T. S., & Kollins, S. H. (2001). Temporal discounting: Basic research and the analysis of socially important behavior. *Journal of Applied Behavior Analysis, 34*(1), 101–122. https://doi.org/10.1901/jaba.2001.34-101

Dijksterhuis, A., Aarts, H. (2010) Goals attention and (Un)consciousness. *Annual Review of Psychology, 61*(1), 467–490. https://doi.org/10.1146/annurev.psych.093008.100445

Dougherty, T. (2015). *Future bias and practical reason*. Michigan Publishing, University of Michigan Library.
Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology, 57*(2), 212–228. https://doi.org/10.1037/0022-3514.57.2.212
Frijda, N. H. (1986). *The emotions* (pp. 12, 544). Editions de la Maison des Sciences de l'Homme.
Ghaffari, M., & Fiedler, S. (2018). The power of attention: Using eye gaze to predict other-regarding and moral choices. *Psychological Science, 29*(11), 1878–1889. https://doi.org/10.1177/0956797618799301
Greene, P., Latham, A. J., Holcombe, A., Miller, K., & Norton, J. (2021b). The rationality of near bias towards both future and past events. *Review of Philosophy and Psychology*. https://doi.org/10.1007/s13164-020-00518-1
Greene, P., Latham, A. J., Miller, K., & Norton, J. (2021a). Hedonic and non-hedonic bias toward the future. *Australasian Journal of Philosophy, 99*(1), 148–163.
Greene, P., Latham, A. J., Miller, K., & Norton, J. (2022a). Why are people so darn past biased? *Temporal Asymmetries in Philosophy and Psychology, 139*, 89.
Greene, P., Latham, A. J., Miller, K., & Norton, J. (2022b). Capacity for simulation and mitigation drives hedonic and non-hedonic time biases. *Philosophical Psychology, 35*(2), 226–252.
Greene, P., Latham, A. J., Miller, K., & Norton, J. (2022c). How much do we discount past pleasures? *American Philosophical Quarterly., 59*(4), 367–376.
Greene, P., Latham, A. J., Miller, K., & Norton, J. (2022d). On preferring that overall, things are worse. *Philosophy and Phenomenological Research., 105*(1), 181–194. https://doi.org/10.1111/phpr.12819
Greene, P., & Sullivan, M. (2015). Against time bias. *Ethics, 125*(4), 947–970.
Hare, C. (2013). Time---the emotional asymmetry
Horwich, P. (1987). *Asymmetries in time: Problems in the philosophy of science*. MIT Press.
Hume, D. (1738) A treatise of human nature. Oxford University Press
Janiszewski, C., Kuo, A., & Tavassoli, N. T. (2013). The influence of selective attention and inattention to products on subsequent choice. *Journal of Consumer Research, 39*(6), 1258–1274. https://doi.org/10.1086/668234
Kauppinen, A. (2018). Agency, experience, and future bias. *Thought, 7*(4), 237–245.
Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). Motivated attention: Affect, activation, and action. In Attention and orienting: Sensory and motivational processes (pp. 97–135). Lawrence Erlbaum Associates Publishers.
Latham, A. J., Miller, K., & Norton, J. (2023). Against a normative asymmetry between near- and future-bias. *Synthese, 201*(3), 93. https://doi.org/10.1007/s11229-023-04045-1
Latham, A. J., Miller, K., Norton, J., & Tarsney, C. (2020). Future bias in action: Does the past matter more when you can affect it? *Synthese, 198*(12), 11327–11349.
Latham, A. J., Miller, K., Oh, J., Shpall, S., & Yu, W. (2024). Exploring arbitrariness objections to time-biases. *Journal of the American Philosophical Association*. https://doi.org/10.1017/apa.2023.14
Latham, A. J., Miller, K., Tarsney, C., & Tierney, H. (2021a). Belief in robust temporal passage (probably) does not explain future bias. *Philosophical Studies*. https://doi.org/10.1007/s11098-021-01748-
Latham, A. J., Miller, K., Tarsney, C., & Tierney, H. (2021b). Robust passage phenomenology probably does not explain future bias. *Synthese*. https://doi.org/10.1007/s11229-022-03514-3
Latham, A. J., Oh, J., Miller, K., Shpall, S., & Yu, W. (2023). Exploring arbitrariness objections to time-biases. *Journal of the American Philosophical Association., 5*, 1–7.
Lawless, L., Drichoutis, A., & Nayga, R. (2013). Time preferences and health behaviour: A review. *Agricultural and Food Economics, 1*(1), 1–19.
Lee, R., Hoerl, C., Burns, P., Fernandes, A. S., O'Connor, P. A., & McCormack, T. (2020). Pain in the past and pleasure in the future: the development of past-future preferences for hedonic goods. *Cognitive Science, 44*, e12887. https://doi.org/10.1111/cogs.12887
Lee, R., Shardlow, J., O'Connor, O. H., Hotson, L., Hotson, R., Hoerl, C. & McCormack, T. (2022). Past-future preferences for hedonic goods and the utility of experiential memories. *Philosophical Psychology, 35*(8), 1181–1211. https://doi.org/10.1080/09515089.2022.2038784
Lewis, D. (1979). Counterfactual dependence and times? *Arrows Noûs, 13*(4), 455–476. https://doi.org/10.2307/2215339
Lewis, D. K. (1986). *On the plurality of worlds*. Wiley-Blackwell.
Loewenstein, G., and Jon E., eds. *Choice over time*. Russell Sage Foundation, 1992.

Maclaurin, J., & Dyke, H. (2002). 'Thank goodness that's over': The evolutionary story. *Ratio, 15*(3), 276–292.
Mischel, W., Grusec, J., & Masters, J. C. (1969). Effects of expected delay time on the subjective value of rewards and punishments. *Journal of Personality and Social Psychology, 11*(4), 363.
Mrkva, K., Ramos, J., Van Boven, L. (2020) Attention influences emotion judgment and decision making to explain mental simulation. *Psychology of Consciousness, 7*(4), 404–422. https://doi.org/10.1037/cns0000221
Nagel, T. (1970). Death. *Noûs, 4*(1), 73-80. https://doi.org/10.2307/2214297
Parfit, D. (1984). *Reasons and persons*. OUP Oxford.
Pleskac, T., Yu, S., Grunevski, S., & Liu, T. (2022). Attention Biases preferential choice by enhancing an option's value. *Journal of Experimental Psychology. General*. https://doi.org/10.1037/xge0001307
Price, H. (1994). A neglected route to realism about quantum mechanics. *Mind, 103*(411), 303–336.
Price, H. (1996). Discussion Backward causation and the direction of causal processes: Reply to Dowe. *Mind, 105*(419), 467–474.
Ramos, J., Caruso, E. M., & Boven, L. V. (2022). Temporally asymmetric psychology: prospection, retrospection, and well-being. In C. Hoerl, T. McCormack, & A. Fernandes (Eds.), *Temporal Asymmetries in Philosophy and Psychology.* Oxford University Press.
Rawls, J. (1971). *A theory of justice*. Harvard University Press.
Read, D. (2004). Intertemporal choice. In D. J. Koehler & N. Harvey (Eds.), *Blackwell Handbook of Judgment and Decision Making* (pp. 424–443). Blackwell Publishing.
Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica, 5*, 61–71.
Samuelson, P. A. (1948). Consumption theory in terms of revealed preferences. *Economica, 15*, 243–253.
Samuelson, P. A. (1950). The problem of integrability in utility theory. *Economica, 17*, 355–383.
Sidgwick, H. (1874). Methods of Ethics (E. E. C. Jones, Ed.). Kaplan.
Smith, C., & Lazarus, R. (1990). Emotion and adaptation. *Contemporary Sociology, 21*, 609–637. https://doi.org/10.2307/2075902
Störmer, V. S., & Alvarez, G. A. (2016). Attention alters perceived attractiveness. *Psychological Science, 27*(4), 563–571. https://doi.org/10.1177/0956797616630964
Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences, 30*(3), 299–313. https://doi.org/10.1017/S0140525X07001975
Suhler, C., & Callender, C. (2012). *Thank goodness that argument is over: Explaining the temporal value asymmetry*. Michigan Publishing, University of Michigan Library.
Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters, 8*(3), 201–207.
Yi, R., Gatchalian, K. M., & Bickel, W. K. (2006). Discounting of past outcomes. *Experimental and Clinical Psychopharmacology, 14*(3), 311–317. https://doi.org/10.1037/1064-1297.14.3.311

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Batoul Hodroj[1] · Andrew J. Latham[2] · Kristie Miller[1] · Rasmus Pedersen[1] · Danqi Wang[3]

✉ Kristie Miller
Kristie.miller@sydney.edu.au

Batoul Hodroj
b.hodroj@uq.net.au

Andrew J. Latham
andrew.latham@cas.au.dk

Rasmus Pedersen
rped0728@sydney.edu.au

Danqi Wang
danqi.wang@kcl.ac.uk

1 Department of Philosophy, University of Sydney, Camperdown, Australia

2 Department of Philosophy and History of Ideas, Aarhus University, Aarhus, Denmark

3 Department of Philosophy, Leeds University, Leeds, UK