

7

Caring for Our Future Selves

Alison Fernandes

1. Introduction

Joan is having a dental operation—a painful procedure that must be undertaken without anaesthetic. As she awakes groggily in the hospital, she can't remember whether she has already undergone the surgery, or whether it will take place later that day. She asks a nurse, who goes to find out. As she awaits the reply, would she prefer to learn that the painful experience was in the past or will take place in the future?

If you're like most of us (Caruso, 2018; Lee et al., 2020) and imagine yourself in Joan's situation, you'd prefer your pain was in the past. After all, the past is over and done with—isn't it where you'd prefer your pains and other unpleasant experiences to be? Parfit (1984: Section 64), who first introduced these cases, assumed our preferences would display this temporal asymmetry. Conversely, it seems we have the opposite tendency regarding our pleasures—we prefer pleasures and other pleasant experiences to be in the future, rather than the past, all else being equal (Lee et al., 2020). While it's controversial whether this asymmetry holds for all events, including achievements and disappointments (Parfit, 1984: Section 62; Brink, 2011; Hare, 2013), it seems to hold at least for hedonic experiences—experiences that are pleasant, painful, liked or disliked, in and of themselves. Overall, we seem to have the following *preference asymmetry*: We prefer pains and other unpleasant hedonic experiences to have *already* happened to us in the past (rather than to be in the future) and we prefer pleasures and other pleasant hedonic experiences to be *going to* happen to us in the future (rather than to be in the past), all else being equal.

There are temporal asymmetries in our other attitudes as well. We seem to *value* future events more than past events, in the sense that we tend to judge a given amount of work deserves *more* compensation when it is described as taking place in the future, compared to the past—the 'temporal value asymmetry' (Caruso, Gilbert, and Wilson, 2008; Burns et al., 2019). There also seem to be asymmetries in emotions such as fear, relief, or regret—emotions that may seem appropriate regarding *only* future or *only* past events (Maclaurin and Dyke, 2002).

I fear that the dog *will* bite me (not that he has already) or regret what I *have* done (not what I will do).¹ Call these asymmetries of ‘tensed emotions’.

There are two programs for explaining attitude asymmetries. The first takes attitude asymmetries to reflect deep *metaphysical* differences between the past and future (Prior, 1959; Cockburn, 1998; Craig, 2000: 151–7; Zimmerman, 2007). This program is popular among those defending ‘A-theories of time’. According to an A-theory, the distinctions between past, present, and future events are objective metaphysical differences that are not reducible to person- or perspective-relative facts about when events are located relative to a person or a perspective. The past *in and of itself* is different from the future and has different metaphysical properties. According to this metaphysical program, an event’s being in the past or being in the future explains asymmetries in our attitudes towards it. While we might be related to the event in various ways (it might be in *our* future), our temporal relation to the event is not required to explain temporal features of our attitudes towards it. The temporal properties of the event itself can explain why our attitudes are temporally asymmetric. For example, the fact that we have stronger preferences regarding future events is directly explained by the fact that those events are in the future, rather than the past, where being in the future or past are metaphysical properties that events have in themselves—and that aren’t to be thought of in terms of our relation to them. Call explanations of this form *Temporal Metaphysical Accounts* of attitude asymmetries. If a *Temporal Metaphysical Account* is correct, attitude asymmetries can be used to argue for the truth of an A-theory. If an A-theory is *required* to explain the attitude asymmetries we observe, then the existence of attitude asymmetries supports an A-theory.²

There is an alternative program. This program attempts to explain attitude asymmetries in *scientific* terms, modelling itself after explanatory practice in science and using results from physics, biology, and psychology (Maclaurin and Dyke, 2002; Caruso, Gilbert, and Wilson, 2008; Suhler and Callender, 2012; Prosser, 2016; Callender, 2017). According to this program, attitude asymmetries are to be explained using the particular *relations* between us and events, rather than metaphysical properties that the events have in themselves. For example, the preference asymmetry might be explained by the fact that we have stronger preferences regarding events we can *influence* now—explaining the preference asymmetry in terms of an asymmetry of influence. This program explains attitude asymmetries using features of the world that are countenanced by science and that are acceptable *regardless of the metaphysics of time one adopts*. For this

¹ Sometimes emotions such as regret are *defined* as directed towards past events—see Hoerl and McCormack (2016) and references therein. But I’m inclined to think that these asymmetries are not strict.

² See Fernandes (2019) for more on this argument and why it fails in the case of the value asymmetry. For other criticisms, see Mellor (1998: 41–2), Suhler and Callender (2012), Callender (2017: Ch. 12), and Hoerl (2015).

reason, I will call them *scientific* explanations. While such explanations are compatible with A-theories of time, they are more popular among those defending ‘B-theories’ of time—theories that deny there are objective metaphysical differences between the past, present, and future. If attitude asymmetries can be explained in scientific terms, then there is no explanatory argument from attitude asymmetries to an A-theory of time.

This chapter will support the scientific program for explaining attitude asymmetries. But it will do so by criticising the general form that scientific explanations have taken. The standard scientific explanation is what I will call the *Simple Causal Account*—an exemplar of a certain class of accounts. According to the *Simple Causal Account*, attitude asymmetries are due to a temporal asymmetry in what events our decisions and actions can *influence* now. Because our decisions and actions can potentially influence future events, but cannot even potentially influence past events, there are asymmetries in our preferences (Callender, 2017), value judgements (Caruso, Gilbert, and Wilson, 2008; Suhler and Callender 2012; see also Callender 2017: Ch. 12), tensed emotions (Maclaurin and Dyke, 2002), and attitudes more generally (Prosser, 2016). An asymmetry of influence seems a good place to start to explain attitude asymmetries. Causal explanations are used extensively in the sciences and causation is a robust and pervasive temporally asymmetric relation, at least at our world—causes always come before their effects. It is reasonable to expect our attitudes to be responsive to such an asymmetry.

I will argue, however, that the role of causation in explaining at least some attitude asymmetries has been misunderstood. Some attitude asymmetries are *not* due to an asymmetry in what our actions and decisions can influence now. Instead, they are due to temporal asymmetries in a broader set of causal relations that underlie personal identity and that demarcate a causal ordering in our lives between our ‘past’ and ‘future’ selves. We care more about events that concern our *causally* later selves, where being a causally later self implies a richer set of phenomena beyond mere action-influence, including memory, planning, and anticipation. While asymmetries in these relations are all manifestations of the temporal asymmetry of causation, their contribution to attitude asymmetries is not via what our decisions and actions can influence now—but via how they structure the order of our lives. While these person-connecting relations have been used to explain other temporal phenomena, such as the apparent flow of time (Mellor, 1998: 122–3; Dainton, 2010: Ch. 7; Callender, 2017: Ch. 10; Ismael, 2012; Prosser, 2016: Ch. 7), they have been largely absent from discussions of attitude asymmetries—with the important exception of Gallois (1994). This chapter aims to remedy this defect.

I’ll develop an ‘*Enriched Causal Account*’ that appeals to person-connecting causal relations. This account has a number of advantages. Firstly, it predicts the intuitively correct results in a broader range of cases than both the *Simple Causal Account* and a *Temporal Metaphysical Account*. Secondly, it fits with a broadly

naturalistic and evolutionary approach to explaining attitude asymmetries. Thirdly, it makes interesting testable predictions: it suggests attitude asymmetries will be first-personal and may vary with how connected we feel to our future and past selves. The *Enriched Causal Account* is unlikely to be the full story. But it points to the need to pay more attention to our causal connectedness through time to explain attitude asymmetries.

The *Enriched Causal Account* matters to a number of projects. Firstly, it supports the broad program of explaining temporal asymmetries in scientific terms and undermines the explanatory argument to an A-theory noted above. Secondly, the account matters to empirical work in psychology, suggesting we need to look to how we relate to our past and future selves to understand the development of attitude asymmetries. Thirdly, the *Enriched Causal Account* matters to arguments concerning the *rationality* of attitude asymmetries. The *Enriched Causal Account* suggests attitude asymmetries are inherently perspectival. Insofar as one thinks rational asymmetries in our attitudes *cannot* be perspectival (Sidgwick, 1907; Brink, 2011; Sullivan, 2018: Ch. 7), the *Enriched Causal Account* supports their irrationality. The account also supports arguments against the rationality of attitude asymmetries that rely on facts about personal identity (Greene and Sullivan, 2015; Sullivan, 2018: Ch. 6; Dougherty, 2015). Conversely, the *Enriched Causal Account* specifies on what terms their rationality may be defended—temporal asymmetries are rational insofar as we can justify caring more about what happens to our future selves. Finally, insofar as we wish to ameliorate the effects of attitude asymmetries in cases where we take them to be irrational, the *Enriched Causal Account* suggests how we might do so—by coming to care more equally about our past and future selves.

Before I begin, let me note some features of my methodology. Firstly, while I aim to identify features that may be relevant for explaining other attitude asymmetries, I will focus on the preference asymmetry. The preference asymmetry is the most plausibly strict and rational and is the asymmetry most commonly used to argue *against* a B-theory of time (Cockburn, 1998; Craig, 2000: 151–7; Zimmerman, 2007). If the preference asymmetry can be explained scientifically, there is less hope that a metaphysical temporal explanation is needed to explain more complex attitude asymmetries. See Callender (2017: Ch. 12) and Fernandes (2019) for discussion.

Secondly, I will focus on first-person hedonic prudential cases of the preference asymmetry. These are the cases where the preference asymmetry has been thought to hold most strongly (Parfit, 1984: Section 62; Brink, 2011; Hare, 2013). I will make no assumptions about whether the preference asymmetry holds for all kinds of events, for third-personal cases or for non-prudential cases. Unlike Parfit (1984: Section 69) and Sullivan (2018: 80), I won't assume that the preference asymmetry is absolute—how much we value a past event may depend on features other than its being past.

Thirdly, I will take appropriate methods of eliciting preferences to include asking what would one prefer to *be* the case (a counterfactual), asking what one would prefer to *learn* is the case (a hypothetical), and asking what someone would choose, in a case where they really *could* choose between past and future events (a more complex counterfactual). In the new cases I discuss, I will consult the reader's intuitions. While relying on intuition is obviously not ideal, empirical testing of such cases has not been done. Such testing would also be difficult since subjects may need significant theoretical knowledge to understand the cases. But there may be other ways to test the role of personal identity in attitude asymmetries (section 6).

The chapter proceeds as follows. In section 2, I outline the *Simple Causal Account* of attitude asymmetries. In section 3, I present a counterexample—a case where an attitude asymmetry holds, even though the asymmetry of influence does not. In sections 4–5, I present the *Enriched Causal Account* and argue that it gets the right response in a wider range of cases than alternatives. In section 6, I consider further upshots of the enriched account, concerning its fit with evolutionary explanations and empirical research.

2. The Simple Causal Account

The *Simple Causal Account* explains the preference asymmetry as follows. Assume that, all else being equal, we have stronger preference concerning our own hedonic experiences that we can influence now, compared to those that we cannot. We have stronger preference *towards* events if they are pleasant experiences for us and we can influence them and stronger preferences *against* events if they are unpleasant experiences for us and we can influence them, all else being equal. Having preferences that are asymmetric in this way looks evolutionarily plausible. Our ancestors will have done better to focus their care and attention on events they have some degree of control over (such as whether they will eat tasty food tomorrow), rather than events they cannot control (such as whether they ate tasty food yesterday). An increased focus of care and attention may imply, in turn, having stronger preferences. While I will leave the precise mechanism unspecified, one suggestion is that *emotions* direct our care and attention and that feeling stronger emotions leads to having stronger preferences.³ I will assume an appropriate link can be established between care and attention, on the one hand, and preference.

³ For arguments that emotion plays this role in the case of the value asymmetry, see Caruso, Gilbert, and Wilson (2008), Caruso (2010), Suhler and Callender (2012) and Callender (2017: Ch. 12). For scepticism, see Burns et al. (2019). For more on the emotion asymmetry and the role of emotion in choice, see Newby-Clark and Ross (2003), van Boven and Ashworth (2007) and Gilbert and Wilson (2009).

In the actual world, there is a strict temporal asymmetry of influence: we can influence the past and not the future. According to the *Simple Causal Account*, the distinction between the past and future can therefore function as a heuristic for the distinction between events we can (in principle) influence and those we cannot. The mechanisms governing our preferences ‘overgeneralises’ such that we have stronger preferences towards all future events, even those we cannot (in fact) influence. One suggestion is that differences in how we often simulate future events leads to asymmetries in how we simulate all future events (van Boven and Ashworth, 2007)—see Fernandes (2019) for discussion. So, evolution has selected us to have stronger preferences regarding future events compared to past events.

The *Simple Causal Account* has a number of advantages. Firstly, it fits with evolutionary explanations of behaviour. There are plausible natural mechanisms that can explain how the preference asymmetry arose. Secondly, the account might explain why the preference asymmetry appears rational. The preference asymmetry is the evolutionary upshot of tendencies that have served to promote our overall fitness. Thirdly, the account traces the preference asymmetry back to the temporal asymmetry of causation, an asymmetry accepted by those defending a variety of metaphysical views of time.

Given its advantages, it is no surprise that the *Simple Causal Account* has dominated the scientific program for explaining attitude asymmetries. Early defenders include Hume (2007 [1739–40]: Section 2.3.7.6) and Horwich (1987: 196–8). The recent discussion draws heavily on Maclaurin and Dyke (2002), who defend a *Simple Causal Account* of tensed emotions such as fear, regret, and relief. ‘The special character of past-directed emotions [such as relief] flows from the fact that we are unable to *affect* the states of affairs that are the objects of those emotions’ (at 290, my emphasis). More generally, ‘We care about future pain in a way that we don’t care about past pain because we can avoid future pain’ (ibid. 285).

Since Maclaurin and Dyke (2002), the *Simple Causal Account* has been the go-to scientific explanation for explaining attitude asymmetries. Prosser uses the *Simple Causal Account* to explain attitude asymmetries (2006: 89–90). Caruso, Gilbert, and Wilson (2008) use the simple account to explain the value asymmetry.⁴ Callender (2017), see also Suhler and Callender (2012), uses the *Simple Causal Account* to explain asymmetries of preference and value. Greene and Sullivan (2015: 968), see also Sullivan (2018: Section 8.3), offer a *Simple Causal Account* of emotion and preference asymmetries. While accounts sometimes include other features, such as such as asymmetries of uncertainty (Caruso, 2010; Suhler and Callender, 2012; Callender, 2017: Ch. 12) or the different social roles of permission and punishment (Caruso, 2010: 621), all the accounts just

⁴ In later work, van Boven and Caruso (2015) defend an account of attitude asymmetries based on the ‘temporal doppler effect’—future events feel closer and are cared about more because they are approaching us. The arguments I present apply equally against this account.

mentioned take the *Simple Causal Account* to be correct concerning the contribution of causation to explaining attitude asymmetries.

3. A Shocking Counterexample

Here is a counterexample to the *Simple Causal Account* of preference asymmetry, based on a case by Tarsney (2017).⁵ While Tarsney was concerned with what *justifies* the preference asymmetry, I will be concerned with what *explains* it—though I'll suggest how one could develop a justification below (section 6).⁶ Tarsney took the case to defeat *any* causal account of the preference asymmetry. I'll argue for a different upshot.

Shocking: Joan has had the misfortune to encounter a philosophical fiend. The fiend has subjected Joan to a series of equally painful electric shocks—so many that Joan has lost count and doesn't remember how many she has had. But, after six hours, the fiend's experiment is almost at an end. He now offers Joan the following choice. Either a) Joan will experience five more electric shocks in the next five minutes, or, b) Joan will experience 10 more electric shocks five hours ago. The demon will use his retro-causal device to ensure that the past 10 shocks are brought about if Joan chooses b). While the additional shocks in the past or future would be painful, they make no further difference to Joan's future health and happiness. Assume Joan fully understands the causal structure of her situation.

What should Joan decide? Tarsney (2017) argues that, intuitively, Joan should choose b). Even though this will lead to her experiencing more shocks overall, Joan should prefer to avoid the *future* shock to come, even if she has to have had more shocks in the *past*. After all, the shocks in the past are over and done with. Why should she care if she has a few more?

Assume for the moment that Tarsney is correct—Joan should make choice b). Assume further that agents like Joan *would* choose appropriately in these circumstances. Then there's a problem for the *Simple Causal Account*. The *Simple Causal Account* explains the preference asymmetry in terms of an asymmetry of what the subject can influence *now*. But, in *Shocking*, there is no strict asymmetry of influence. Joan's choices now *can* influence earlier events. If the *Simple Causal Account* were correct, it seems Joan would be equally concerned about the past and future shocks. So she would choose a)—which implies experiencing less shocks overall. But it seems that Joan would choose b). So the

⁵ Parfit (1984: Section 65) argues against a control-based justification of attitude asymmetries by arguing that even if future torture was outside our control, we wouldn't care about it less. The response I make to *Shocking* applies equally to this argument.

⁶ Some of defenders of the *Simple Causal Account* explicitly take themselves *not* to be attempting to justify the asymmetry (Horwich, 1987: 196; Callender, 2017: 289) including some Tarsney (2017: 765, fn. 6) identifies as doing so (Suhler and Callender, 2012: 6).

simple account makes the wrong prediction. If the preference asymmetry holds even when the influence asymmetry does not, it seems the influence asymmetry cannot explain the preference asymmetry.

Here are three possible responses. Firstly, one might reject the possibility of backwards causation, even in hypothetical contexts, and so reject the relevance of *Shocking*. Dialectically, however, this is not a strong move. Defenders of the *Simple Causal Account* aim to show that attitude asymmetries can be explained using a temporal asymmetry of causation, rather than an asymmetry in time itself. Cases of backwards causation are one of the most effective means of disentangling these potential contributions (Fernandes 2020, 2021). Defenders of the *Simple Causal Account* should allow that cases like *Shocking* are possible.

Secondly, one might reject the claim that Joan *would* make choice b). Perhaps Joan would choose less shocks overall, no matter their temporal location. Recent empirical work suggests that the preference asymmetry is not strict, even for real-world cases. Most people report preferring future pain, if there is twice as much past pain as future pain (Caruso, 2018; Lee et al., 2020). Any degree of preference asymmetry, however, is enough to generate the *Shocking* counterexample, since any difference of degree cannot be accounted for by a difference in what we can influence. While I'll argue (section 5) that there are cases importantly similar to *Shocking* where the Joan would intuitively choose a), absent direct empirical research, I'll allow that Tarsney is correct and that Joan would choose b).

A third response is to explain why Joan would choose b), while still retaining the *Simple Causal Account*. For example, one might argue that because Joan has been evolutionarily hard-wired to direct more care and attention towards future events, her preference asymmetry will not be disrupted by a one-off case in which she can influence the past. For example, say van Boven and Ashworth (2007) are correct that we experience stronger emotions towards future events because of framing effects—we simulate future events in certain characteristic ways and so experience stronger emotions concerning them. These stronger emotions may lead us to have stronger preferences regarding future events, even in cases like *Shocking*. Call this the *temporal-heuristic response*.

The account I will defend is compatible with aspects of the *temporal-heuristic response*. I agree that a heuristic we adopt in the actual world can account for why Joan would choose b) in a case like *Shocking*. However, in the next section I will argue that the heuristic we adopt is *not* a simple past/future framing. Causation's contribution to the preference asymmetry is more complex than the *Simple Causal Account* suggests.

4. The Enriched Causal Account

In order to make room for an alternative causal account, we should consider in what sense the past shocks might seem 'over and done with' in *Shocking*. One

suggestion, following A-theorists, is that events are ‘over and done with’ if they are in the temporal past. Another suggestion, following the *Simple Causal Account*, is that events are ‘over and done with’ if they can’t be influenced now. But a third alternative is that events are ‘over and done with,’ for you, when they are in *your* past, where what is in *your* past relates to what you have already lived through and what has been incorporated into your life history up to now—which may be distinct from what has happened in the past or what you cannot influence.

In order to develop this idea, I will adopt a distinction from the literature on time travel: ‘external time’ versus ‘personal time’ (Lewis, 1976a). *External* time is what is measured by clocks and other mechanisms in the actual world and in parts of time travel worlds that don’t involve time travel. For example, as an observer, you might see Dr. Who depart in her time machine at 12 p.m. and arrive back at 1 p.m.—one hour has passed in external time. *Personal* time, by contrast, is measured by clocks and other mechanisms in the vicinity of *particular people*—including time travellers. For example, after Dr. Who departs in her time machine at 12 p.m., she might spend one week visiting distant parts of the future and return (at 1 p.m.) *having aged one week*. So one week has passed in her personal time. In the actual world, without time travel, personal time and external time always align. But, in time travel worlds, they can come apart. In the above case, only one hour passes between Dr. Who’s departure and arrival in external time, while one week passes in her personal time. Note that while personal time is typically discussed with respect to people, it remains an objective measure of time—there are facts about how much personal time has passed, facts we can be wrong about.

What determines facts about personal time (in cases involving people) are ultimately the causal relations that order our lives in the actual world. Consider Dr. Who’s travels. The reason why one week passes in her personal time as she travels is because she *ages* one week—she is physically a week’s older-looking, has built up a week’s worth of memories and has undergone physical and psychological changes that would, for a non-time traveller, usually take a week of external time. So she is one week older in personal time.

The same causal relations that are used to demarcate personal time are also used in determining *personal identity*—what makes someone the very same person over time. More precisely, these personal-identity conferring causal relations determine what makes two ‘person-segments’ at different times (or places) actually parts of the very same person.⁷ Consider, for example, the most popular contemporary accounts of personal identity—psychological-continuity accounts. According to psychological-continuity accounts, what makes two person-segments parts of the same person is the fact that there are psychological similarities between the two segments or similarities between a series of person-segments connecting

⁷ I intend talk of ‘person-segments’ to be neutral regarding the metaphysics of persistence.

them—implying psychological *continuity* between the two segments. One segment might have most of the memories of the earlier segment, for example. But, as Lewis (1976b: 17), Shoemaker (1984: 90) and Parfit (1984: Section 207) point out, it can't be mere *similarity*-based continuity that makes these person-segments parts of the same person, since similarity may be brought about by artificial or coincidental means—I might engineer someone to have your memories, even though they are not you. So, according to psychological-continuity accounts, for the similarity-based continuity to contribute to personal identity, it has to be brought about by psychological states of earlier segments *appropriately causing* the psychological states of directly later segments.

Similar causal relations also feature in the other major contemporary accounts of personal identity. For example, according to physical-continuity theories, personal identity is brought about by two person-segments being the same biological animal (Williams, 1970). But being the same biological animal is, again, underwritten by causal relations that connect the physical animal-segments across time. According to narrative theories, personal identity is brought about, roughly, by a person building up a coherent narrative history that identifies certain person-segments as earlier segments of themselves (Schechtman, 1996: Ch. 5). But, once again, what allows someone the resources to build up such narratives are causal mechanisms—mechanisms such as memory that grant us relatively direct epistemic access to experiences of causally earlier segments, which will turn out to be the experiences of earlier *selves* in the narratives we develop.

In both discussions of time travel and personal identity, causal relations are used to induce an ordering on a person's life. For this reason, I'll use '*causally earlier selves*' and '*causally later selves*' to refer to earlier and later segments of a person's life as given by their causal ordering. This talk involves no worrying circularity. When I speak of memories of a causally earlier self, for example, what *makes* that segment a causally earlier self is precisely that the relevant causal relations hold between this segment and a causally later segment.

I won't attempt a precise account of what causal relations order our lives—what 'appropriate' causal relations there must be for selves to count as segments of a single person as well as what degree of physical or mental continuity or narrative coherence is required. But here are some of the phenomena that are likely to be important:

- *Memory and anticipation*: We remember experiences of causally earlier selves, remember remembering experiences of even causally earlier selves and anticipate causally future selves remembering experiences of earlier selves (Mellor, 1998: 122–3). Moreover, the memories of our causally later selves are caused, via somewhat stable mechanisms, by the experiences and memories of our causally earlier selves.
- *Beliefs, desires, character, and dispositions*: The beliefs, desires, character, and dispositions of our causally later selves are caused in part by the beliefs,

desires, character, and dispositions of our causally earlier selves, as well as other states, interactions, and experiences that our causally earlier selves undergo.

- *Plans and intentions*: The beliefs, desires, plans, intentions, and actions of our causally later selves are caused in part by the plans and intentions of our causally earlier selves, allowing us to engage in extended actions and projects.
- *Emotions*: The emotions of our causally later selves are caused by the emotions of our causally earlier selves, as well as by the character of our causally earlier selves and other states, interactions, and experiences of our causally earlier selves.
- *Ageing*: The physical constitution of our causally later selves is caused by the physical constitution of our causally earlier selves, as well as by other physical states and interactions of our causally earlier selves.

The appropriate causal relations that underwrite these phenomena hold *directly* between ‘causally adjacent selves’—person segments of the same self whose causal relations to each other are not typically mediated by other person segments. Causally earlier and later selves are connected by a chain of causally adjacent selves, so that the appropriate causal relations may only hold indirectly between them.

I haven’t specified exactly how closely related causally adjacent or non-adjacent selves need to be to count as selves of the same person—what precise continuity or coherence criteria are required for a full account of personal identity. Nor have I said anything about how to handle cases where criteria might conflict or are apparently satisfied by ‘too many’ selves. But such concerns are no reason to reject the general idea that particular causally ordered phenomena underlie personal identity: differences of degree and the potential for conflict and other problem cases are what we should expect for a higher-level and socially relevant kind such as ‘personhood’. Nor will my arguments depend on any particular way of resolving problem cases.

What will be important for the *Enriched Causal Account* is that there are particular phenomena that *causally* order our lives and that these causal phenomena are distinct from relations of mere influence. Standard accounts of personal identity all rely on particular causal phenomena to bring about the continuity and coherence required to be the same self over time. These phenomena are causally ordered, but, importantly, they have other features as well—they concern particular *kinds* of states (memories, physical states, etc.) related in particular ways and that are *stable* to a certain extent. This stability is needed to secure a sufficient degree of continuity and coherence across selves of the same person. With respect to memory, for example, we expect later selves to ‘retain’ the memories of earlier selves, bringing about continuity in the memories selves have over time. There are ways for us to influence other selves, but these lack the stability or other particular

features that would imply we are parts of the same person. For example, you might convince your friend to come holidaying with you, affect the contents of her memories of the Greek islands, and be responsible for her sunburn. But none of these forms of influence tempts us to suggest that she is in fact a later causal part of you. The reason for this is that these relations of mere influence don't secure sufficient continuity or coherence in causally adjacent selves with respect to the particular phenomena that underlie personal identity.

As a result, it is not just any causal relations between an agent now and *even events that concern that very agent* that establishes the causal order of her life. Even relations of influence *within a single agent* are not equivalent to the causal ordering of a life. Consider the following case. Dot, aged 23, goes back in time and convinces her 8-year-old self to follow her passion and study science. This conversation has a profound effect on young-Dot's life. In this case, adult-Dot *influences* child-Dot, including influencing her beliefs, desires, and plans. But this does not mean that child-Dot comes *after* adult-Dot in the causal ordering of Dot's life; there aren't the appropriate causal dependencies of child-Dot on adult-Dot (via a series of causally adjacent selves). For example, while child-Dot's memories *cause* some of adult-Dot's memories via the usual mechanisms whereby memories are retained, adult-Dot's memories don't cause child-Dot's memories via such mechanisms. While adult-Dot does causally influence child-Dot's beliefs, desires, and plans, the connection here does not go via the specific phenomena identified above that causally order our lives. Adult-Dot has to *convince* child-Dot and can't rely on the physical and psychological mechanisms that make us the same person over time. Even though child-Dot and adult-Dot are causally connected and are the same person, and even if adult-Dot influences child-Dot, adult-Dot still comes *after* child-Dot in the causal order of Dot's life.

According to the *Enriched Causal Account* the causal ordering of our lives can be used to explain preference asymmetries. Begin with the claim that, all things being equal, we prefer pleasant events to happen to our causally later selves (rather than our causally earlier selves) and we prefer unpleasant events to happen to our causally earlier selves (rather than our causally later selves). In the actual world, causal order and temporal order align—events that are causally later are always temporally later. So the causal ordering of our lives also aligns with temporal order. Having asymmetric preferences concerning our causally past and future selves will lead to us having temporally asymmetric preferences: we will prefer pleasant events to happen to us in the *future* rather than the past and unpleasant events to happen to us in the *past* rather than the future.

So far so good. But, to distinguish between the *Enriched Causal Account* and other accounts of attitude asymmetries, we should consider cases where temporal and causal order don't align—cases of backwards causation. In *Shocking*, Joan can influence events that are temporally earlier. So causal order does not align with temporal order. The intuitive prediction was that Joan would care about the future

event more than the past event, even though she can influence the past event. The *Simple Causal Account* could not, therefore, account for her preferences. But the *Enriched Causal Account* can. What happens to Joan at an earlier time is still something that happens to her *causally earlier self*—even though she can influence these events. Recall, mere influence does not change the causal order of our lives. Because Joan prefers unpleasant things to happen to her causally earlier self (rather than her causally later self), she will prefer unpleasant events to happen in the past rather than the future in *Shocking*. So she will make choice b).

Overall, *Shocking* is not a counterexample to any causal account. It is only a counterexample to an account that takes the relevance of causation to go entirely via what we can influence now—not to an account concerned with the much richer ways in which causal relations structure our lives.

5. Time Travel Shocks

There is a second kind of case that can distinguish the *Enriched Causal Account* from other accounts—cases of backwards time travel. Gallois (1994) presents one such case. Gallois' case is briefly discussed in Dainton (2010: 37), but otherwise time travel cases are largely absent from discussions of attitude asymmetries. This is a pity. They are important for disentangling causal and temporal explanations. I'll argue that a second kind of case, *Time Travel Shocks*, gives us reason to reject a *Temporal Metaphysical Account* of attitude asymmetries as well as the modified *Simple Causal Account* that uses the *temporal heuristic response*. Here is the case:

Time Travel Shocks: Jamal has also had the misfortune to encounter a philosophical fiend. The fiend has subjected Jamal to a series of equally painful electric shocks—so many that Jamal has lost count and doesn't remember how many he has had. But, after six hours, the fiend's experiment is almost at an end. He now offers Jamal the following choice. Either a) Jamal will experience five more electric shocks (in the next five minutes), or, b) Jamal will experience 10 more electric shocks five hours ago—and the fiend will send Jamal back in time by five hours using his time machine so that Jamal can live through and experience these past shocks. While the additional shocks in the past or future would be painful, they make no further difference to Jamal's future health and happiness. Assume Jamal fully understands the causal structure of his situation.

What will Jamal decide? I claim that, insofar as Joan will choose the past shocks in *Shocking* (choice b)), Jamal will choose the future shocks in *Time Travel Shocks* (choice a)). Because Jamal will travel back in time in order to experience the past shocks, they are, from Jamal's point of view, very much still to come and promise to be just as painful as the shocks he has experienced so far. These past shocks aren't over and done with. Given Jamal has to choose between shocks that are both still to come from his point of view, he will choose less shocks overall—choice a).

The *Enriched Causal Account* gets this result. According to the *Enriched Causal Account*, Jamal has stronger preferences regarding what happens to his causally future self. In *Time Travel Shocks*, both choices concern his causally future self. So, he will choose less shocks overall—choice a).

Accounts that explain attitude asymmetries merely by appeal to temporal features of events, however, do not imply this prediction. According to a *Temporal Metaphysical Account*, we have significantly weaker preferences concerning events in the past, compared to equivalent events in the future—explaining why Joan prefers 10 past shocks to five future shocks and makes choice b). But this account implies that Jamal would prefer 10 past shocks to five future shocks and would make choice b). Causal features of the setup, including Jamal’s travels in time, should have no effect on his preferences, according to a *Temporal Metaphysical Account*, since they don’t change the temporal structure of Jamal’s situation. But we would expect Jamal to make choice a). A *Temporal Metaphysical Account* makes the wrong prediction.

Time Travel Shocks is not a counterexample to the *Simple Causal Account*, in its unmodified version—since both the past and future shocks are under Jamal’s influence, he will choose less shocks overall. But *Time Travel Shocks* is a counterexample to a *Simple Causal Account* that uses the *temporal heuristic response* to get the intuitive result in *Shocking*. Insofar as framing effects would lead Joan to prefer the past shocks to the future shocks in *Shocking*, we would expect those same effects to lead Jamal to prefer the past shocks to the future shocks in *Time Travel Shocks*. But this is not the intuitive result.

Overall, only the *Enriched Causal Account* recovers the intuitive results in both *Shocking* and *Time Travel Shocks*, as well as in the actual world (see Figure 7.1). While other accounts get the right result in some cases, they cannot explain why

Account	Actual World	<i>Shocking</i>	<i>Time Travel Shocks</i>
<i>Temporal Metaphysical Account</i>	✓	✓	×
<i>Simple Causal Account</i>	✓	×	✓
<i>Simple Causal Account with Temporal Heuristic</i>	✓	✓	×
<i>Enriched Causal Account</i>	✓	✓	✓

Figure 7.1. The success of different accounts of temporal asymmetries in explaining our intuitive preferences. Only the *Enriched Causal Account* delivers the intuitive result in all three kinds of cases.

an agent's preferences will switch between *Shocking* and *Time Travel Shocks*, since the temporal location of the shocks, or whether they can be influenced by the agent now, does not change between *Shocking* and *Time Travel Shocks*. But the *Enriched Causal Account* can explain the switch. In *Time Travel Shocks*, but not *Shocking*, the past shocks are experienced by a causally later self. The *Enriched Causal Account* is best placed to explain the preference asymmetry, since it is the only account that can deliver the intuitive results in all three cases.

6. Further Work

In this final section, I consider an objection to the *Enriched Causal Account*. Responding to this objection will lead to an exploration of how the account fits with evolutionary explanations and empirical work on attitude asymmetries.

Here's the objection. One might worry that the *Enriched Causal Account* needs to explain *why* we care more about our causally later selves. But the reason why we care more about our future selves is simply because we can influence them. So the *Enriched Causal Account* collapses to a *Simple Causal Account*.

Here is an alternative explanation, one that does not collapse the two accounts. The reason why we care more about what happens to our causally later selves is that acting in specific ways towards our causally later selves (in ways that reflect having stronger future-self-directed attitudes) improves our fitness and increases the chance of our genes being inherited.⁸ Here are two kinds of relations that make possible these future-self-directed attitudes and contribute to explaining why these attitudes are so important (see also section 4). Firstly, there are strong *epistemic* relations linking single organisms across time. We acquire information about what happens to ourselves and our immediate surrounds and retain this information into the future by mechanisms such as memory. While we have knowledge of states beyond ourselves, our access to what happens to ourselves is typically much more reliable. Secondly, the causal relations that link organisms across time are much more stable and pervasive than those that link the organism to its environment. While we can influence our surrounds, and other people, we typically have a stronger degree of influence on and more ways to influence our causally later selves. For example, even when we can't influence events in our surrounds, we can still always causally influence our future selves, by controlling how we respond to events and how we allow them to affect our future behaviour. Mechanisms such as these put us in a good position, causally and epistemically, to manage our future behaviour, which ultimately improves our fitness. See Ismael (2016) for further discussion of the mechanisms involved.

⁸ I remain neutral on how precisely having an attitude relates to the manifestation of the behaviour typically associated with that attitude.

There will be some cases where our fitness, or the chances of our genes being inherited, will be enhanced by caring about our causally past selves, about other selves (particularly kin and those in close social communities), and about future events that do not concern our causally future selves. But, given our epistemic and causal situation, this will not be the norm. A creature that is reliable at enhancing its fitness will typically need to direct more of its attention towards managing its own causally future selves and so will care about them more. Its attitudes including emotion and preferences will reflect this greater concern. So we will evolve to care more about and feel attached towards our future selves. These forms of future-self-directed attitudes can build up slowly through evolution.

Overall, the preference asymmetry is partly due to causal asymmetry. It is because we can influence our future selves that we care more about our future selves *compared to our past selves*. So the *Simple Causal Account* gets something right. But the preference asymmetry is also due to our ways of engaging with the world as extended selves and the fact that improving our fitness requires caring more about what happens to *our* future selves, rather than simply what happens in general. This is something the *Simple Causal Account* misses.

So far I have focussed on the preference asymmetry. But the *Enriched Causal Account* can plausibly explain other attitude asymmetries as well, such as those of tensed emotion and value. For example, the fact that we fear some future events (but hardly any past events) might be explained by the fact that we fear what, to our knowledge, may harm our *causally later selves*—and unknown harms to our causally later selves usually come from future events. The fact that we expect more reward or compensation now for future events compared to past events might be explained by the fact that we expect more compensation for what happens to our causally later selves, compared to what happens to our causally earlier selves. The *Enriched Causal Account* offers a basis for explaining a range of attitude asymmetries.

The *Enriched Causal Account* can also explain attitude asymmetries that appear third-personally, such as judging other people to have more effective wills in the future (Helzer and Gilovich, 2012). It is plausible that the same kinds of mechanisms that have led us to behave in ways that are self-directed, also lead us to identify other people as leading causally ordered lives, such that we come to care more about what happens to *their* causally future selves (rather than what happens in general). For example, Hare (2008) argues, in the case of the preference asymmetry, that we should adopt another's perspective on events. Similarly, we might judge other people to have more effective wills in the future, not because we think about the future in general in optimistic terms (Helzer and Gilovich, 2012), but because we think of other people as living causally ordered lives.

A care for future selves may even become generalised to such an extent that it *does* manifest as a care for the future in general (a temporal heuristic), given people's causal futures are always in the temporal future in the actual world—see Figure 7.2. The temporal framing effects I discussed above (section 2) may even

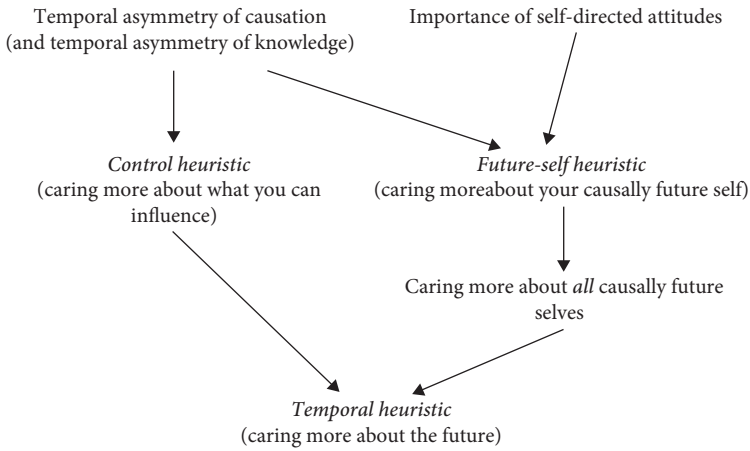


Figure 7.2. Two potential explanatory pathways for explaining attitude asymmetries. Arrows indicate directions of explanation.

contribute to this effect. But their contribution would go via a care for our future selves and a *'future-self heuristic'*.

It is, of course, up to empirical testing to determine what explains an attitude asymmetry in any given case. I suspect *'control heuristics'* of the kind the *Simple Causal Account* appeals to and *temporal heuristics* have both shaped our attitude asymmetries. For example, I expect that temporal framing effects (van Boven and Ashworth, 2007; Wilson and Gilbert, 2008; Caruso, 2010; van Boven, Kane, and McGraw, 2010; Kane et al., 2012) have an important role to play in explaining asymmetries of value, particularly value asymmetries in third-person cases (Caruso, Gilbert, and Wilson, 2008). Similar kinds of framing effects may also be needed to explain why we bet more optimistically on the future than the past (Strickland, Lewicke, and Katz, 1966), why we judge future actions as more due to the will (Helzer and Gilovich, 2012), why we judge past transgressions less harshly (Caruso, 2010), and why we judge that future events feel closer (Caruso et al., 2013).

I suspect that the *future-self heuristic* (with no temporal addition) may play a greater role in explaining attitude asymmetries that appear rational or justified, while *temporal heuristics* may play a greater role to play in explaining attitude asymmetries that appear irrational or unjustified. Caring more about our future selves appears rational, while caring more about the *future* (in and of itself) does not—leading to a difference in how the asymmetries they lead to are judged. But further studies would be needed to test this hypothesis.⁹ Rationally endorsed asymmetries may include the preference asymmetry concerning first-person hedonic goods (Lee et al., 2020; Caruso, 2018; Greene et al., 2021) and an

⁹ See Sullivan (this volume) for skepticism on this point.

asymmetry in how effective the will is taken to be (Helzer and Gilovich, 2012)—at least to the extent that these asymmetries remain when past and future events are compared. Asymmetries that are not rationally endorsed in this way include the value asymmetry (Caruso, Gilbert, and Wilson, 2008).

Other questions the *Enriched Causal Account* does not settle include how manipulatable or high-level the *future-self heuristic* is. It may be that by altering how connected we feel to our causal selves, or how strongly we identify ourselves with them, we can manipulate the preference asymmetry. For example, it may be that we view our past selves from an observer-like perspective (Prinin and Ross, 2006)—see also Buehler, Griffin, and Ross (1994), Trope and Liberman (2010), and O'Brien (2015). Encouraging ourselves to think of our future selves from an observer perspective, or, conversely, to identify more strongly with our past selves, may lead to reductions in the preference asymmetry. This is one way in which the *Enriched Causal Account* might be tested.

There is some limited support for the *Enriched Causal Account* from studies of the preference asymmetry. The *Enriched Causal Accounts* predicts that attitude asymmetries may be perspectival, at least to a degree—that is, stronger in first-person cases compared to third-person cases. Some recent studies suggest that the preference asymmetry is perspectival, to a degree. While 89% of participants preferred painful past surgery to equally painful future surgery for themselves, only 65% have this preference for other people (Caruso, 2018). Similarly, preferences for four hours of past pain over two hours of future pain drop from 61% (first-person) to 44–45% (third-person) (ibid.). Related drops are reported for preferences and emotions concerning future pleasures (Lee et al., 2020) and other hedonic goods (Greene et al., 2021).¹⁰ While an attitude asymmetry being perspectival does not distinguish the *Enriched Causal Account* from the *Simple Causal Account*, it does suggest that temporal heuristics on their own aren't enough to account for that attitude asymmetry.

The *Enriched Causal Account* also has interesting upshots for how we measure the preference asymmetry and other attitude asymmetries. The standard methodology, following Parfit (1984: Section 64), is to present a hypothetical case in which a past or future painful surgery is followed by a period of amnesia. This amnesia is introduced to ensure: a) that the past surgery has no further influence on one's experiences now, such as by producing painful memories; and b) that you don't *know* whether the past surgery has occurred, and so plausibly can have

¹⁰ In Lee et al. (2020) 92% of adults preferred a pleasurable event to be in the future rather than the past, a rate above chance, but only 67% thought that others would be happier in this case, a rate not significantly above chance (Experiments 1a, 1b). 82% thought that others would prefer a pleasurable event to be in the future (Experiment 2), but this was not significantly different to the proportion of adults who reported having this preference for themselves. Perspectival effects were not recorded for preferences concerning painful events. Note that while Caruso (2018) and Greene et al. (2021) asked what one would *prefer* to be the case for someone else, Lee et al. asked what one would *expect* to be the case concerning another's preferences or emotions.

preferences concerning what you will learn. If the *Enriched Causal Account* is right, the period of amnesia introduces a potential confounder. It disrupts the epistemic and causal structures that underlie personal identity and prevents us identifying as strongly with our past selves. Even if we *did* care equally about our past and future selves in general, we might be led to prefer the past pain in the surgery case merely because the amnesia means that the past pain can be treated as pain that *happens to someone else* and that isn't part of one's causal life history. The future pain, even if it is followed by a period of amnesia, is not disconnected in the same way from one's causal history *now*—it is to be treated as something that will *happen to me*. This potential confounder is a concern both for the empirical task of measuring preference asymmetries, as well as for philosophical arguments that appeal to intuitions about the rationality of these asymmetries (Parfit 1984: Section 64). The concern is particularly strong for arguments, like Parfit's, that use intuitions about the rationality of these asymmetries on the way to arguing that personal identity facts are not important to what we ought to do. The relations underlying person identity can't be assumed to be not important, on the way to arguing that they're not.

There are alternatives to the amnesia case. When testing the preference asymmetry, Lee et al. (2020) asked people who they would prefer *to be* (rather than what they would prefer to *learn* is the case) and specified that the people they might be weren't currently experiencing pain or pleasure now due to the event. Even young children could understand these counterfactuals. Another alternative is to use cases where the amnesia is local (concerning only the pain), temporary, and/or accidental in order to minimise the extent to which the amnesia disrupts the relations underlying personal identity. In *Shocking*, and in Tarsney (2017), Joan is described as having experienced so many shocks she can't remember how many she has experienced—this forgetting is local, accidental, and not obviously part of the setup.

7. Conclusion

The *Simple Causal Account* cannot explain attitude asymmetries such as the preference asymmetry because it relies merely on an asymmetry in what we can influence now. Instead, we need the *Enriched Causal Account*: an account that appeals to the much richer set of relations that underlie the causal ordering of our lives.

The *Simple Causal Account* does look tempting. It is simple and seems to capture important aspects of the A-theorist's explanation of attitude asymmetries. The A-theorist uses a global temporal order to explain attitude asymmetries. The B-theorist cannot do this, but they can look for a global temporally asymmetric *relation* that stands in for the A-theorist's global temporal asymmetry—and causation seems to be the right kind of relation to use. Causal relations divide the world into what comes causally after and what comes causally before.

But this move oversimplifies the explanation of attitude asymmetries. The B-theorist should not aim to ape the A-theorist's explanation. To do so misses the more nuanced ways in which our perspective and the causal order of our lives affect how we engage with the world. There has been an increasing interest in considering the structure and perspective of agents in giving scientific explanations of temporal asymmetries and time's apparent features (Mellor, 1998: 122–3; Price, 2007; Dainton, 2010: Ch. 7; Ismael, 2012; Callender, 2017; Prosser, 2016: Ch. 7; Fernandes, 2017). It is time to apply this lesson to the case of attitude asymmetries as well.¹¹

References

- Brink, D. O. (2011). Prospects for temporal neutrality. In C. Callender (Ed.), *The Oxford Handbook of Time* (pp. 353–81). Oxford: Oxford University Press.
- Buehler, R., Griffin, D., and Ross, M. (1994). Exploring the 'planning fallacy': Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67, 366–81.
- Burns, P., McCormack, T., Jaroslawska, A., Fitzpatrick, A., McGourty, J., and Caruso, E. (2019). The development of asymmetries in past and future thinking. *Journal of Experimental Psychology: General*, 148(2), 272–88.
- Callender, C. (2017). *What Makes Time Special*. Oxford: Oxford University Press.
- Caruso, E. M., Gilbert, D. T., and Wilson, T. D. (2008). A wrinkle in time: Asymmetric valuation of past and future events. *Psychological Science*, 19(8), 796–801.
- Caruso, E. M., Van Boven, L., Chin, M., and Ward, A. (2013). The temporal Doppler effect: When the future feels closer than the past. *Psychological Science*, 24(4), 530–6.
- Caruso, E. M. (2010). When the future feels worse than the past: A temporal inconsistency in moral judgment. *Journal of Experimental Psychology: General*, 139(4), 610–24.
- Caruso, E. M. (2018). Why the future is bigger (and 'badder') than the past. Presentation at *Temporal Asymmetries in Philosophy and Psychology Workshop*, University of Warwick.
- Cockburn, D. (1998). Tense and emotion. In R. Le Poidevin (Ed.), *Questions of Time and Tense* (pp. 77–91). Oxford: Clarendon Press.
- Craig, W. (2000). *The Tensed Theory of Time*. Dordrecht: Kluwer Academic Publishers.
- Dainton, B. (2010). *Time and Space*. 2nd ed. Montreal: McGill Press.
- Dougherty, T. (2015). Future-bias and practical reason. *Philosophers' Imprint*, 15(30), 1–16.

¹¹ My warm thanks to Christoph Hoerl, Alastair Wilson, Ruth Lee, Teresa McCormack, Patrick Burns, and Craig Callender for helpful comments and suggestions.

- Fernandes, A. (2017). A deliberative approach to causation. *Philosophy and Phenomenological Research*, 95(3), 686–708.
- Fernandes, A. (2019). Does the temporal asymmetry of value support a tensed metaphysics? *Synthese*, 198(5), 3999–4016.
- Fernandes, A. (2020). Freedom, self-prediction, and the possibility of time travel. *Philosophical Studies*, 177(1), 89–108.
- Fernandes, A. (2021). Time travel and counterfactual asymmetry. *Synthese*, 198 (3), 1983–2001.
- Gallois, A. (1994). Asymmetry in attitudes and the nature of time. *Philosophical Studies*, 76(1), 51–69.
- Gilbert, D. T., and Wilson, T. D. (2009). Why the brain talks to itself: Sources of error in emotional prediction. *Philosophical Transactions of the Royal Society B*, 364, 1335–41.
- Greene, P., and Sullivan, M. (2015). Against time bias. *Ethics*, 125, 947–70.
- Greene, P., Latham, A., Miller, K., and Norton, J. (2021). Hedonic and non-hedonic bias towards the future. *Australasian Journal of Philosophy*, 99(1), 148–63.
- Hare, C. (2008). A puzzle about other-directed time-bias. *Australasian Journal of Philosophy*, 86(2), 269–77.
- Hare, C. (2013). Time—the emotional asymmetry. In H. Dyke and A. Bardon (Eds.), *A Companion to the Philosophy of Time* (pp. 507–20). Chichester: John Wiley & Sons.
- Helzer, E. G., and Gilovich, T. (2012). Whatever is willed will be: A temporal asymmetry in attributions to will. *Personality and Social Psychology Bulletin*, 38(10), 1235–46.
- Hoerl, C. (2015). Tense and the psychology of relief. *Topoi*, 34(1), 217–31.
- Hoerl, C., and McCormack, T. (2016). Making decisions about the future: Regret and the cognitive function of episodic memory. In K. Michaelian, S. Klein, and K. Szpunar (Eds.), *Seeing the Future: Theoretical Perspectives on Future-Oriented Mental Time Travel* (pp. 241–66). New York: Oxford University Press.
- Horwich, P. (1987). *Asymmetries in Time: Problems in the Philosophy of Science*. Cambridge, MA: MIT Press.
- Hume, D. (2007). [1739–40] *A Treatise of Human Nature*. D. Fate Norton and M. J. Norton (Eds.) Oxford: Clarendon Press.
- Ismael, J. (2012). Decision and the Open future. In A. Bardon (Ed.), *The Future of the Philosophy of Time* (pp. 149–68). London: Routledge.
- Ismael, J. (2016). *How Physics Makes Us Free*. New York: Oxford University Press.
- Kane, J., Van Boven, L., and McGraw, A. P. (2012). Prototypical prospection: future events are more prototypically represented and simulated than past events. *European Journal of Social Psychology*, 42(3), 354–62.
- Lee, R., Hoerl, C., Burns, P., Fernandes, A., O'Connor, P. A., and McCormack, T. (2020). Pain in the past and pleasure in the future: Past-future preferences for hedonic goods. *Cognitive Science*, 44(9).

- Lewis, D. (1976a). The paradoxes of time travel. *American Philosophical Quarterly*, 13(2), 145–52.
- Lewis, D. (1976b). Survival and identity. In A. Rorty (Ed.), *The Identities of Persons* (pp. 17–40). Berkeley: University of California Press. Reprinted in Lewis, D. *Philosophical Papers* vol. I, New York: Oxford University Press, 1983.
- Maclaurin, J. and Dyke, H. (2002). ‘Thank goodness that’s over’: The evolutionary story. *Ratio*, 15(3), 276–92.
- Mellor, D. H. (1998). *Real Time II*. London: Routledge.
- Newby-Clark, I. R., and Ross, M. (2003). Conceiving the past and future. *Personality and Social Psychology Bulletin*, 20, 807–18.
- O’Brien, E. (2015). Mapping out past and future minds: The perceived trajectory of rationality versus emotionality over time. *Journal of Experimental Psychology: General*, 144(3), 624–38.
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Clarendon Press.
- Price, H. (2007). Causal perspectivalism. In H. Price and R. Corry (Eds.), *Causation, Physics, and the Constitution of Reality: Russell’s Republic Revisited* (pp. 250–92). Oxford: Oxford University Press.
- Prior, A. N. (1959). Thank goodness that’s over. *Philosophy*, 34(128), 12–17.
- Pronin, E., and Ross, L. (2006). Temporal differences in trait self-ascription: When the self is seen as an other. *Journal of Personality and Social Psychology*, 90, 197–209.
- Prosser, S. (2006). Temporal metaphysics in Z-land. *Synthese*, 149(1), 77–96.
- Prosser, S. (2016). *Experiencing Time*. Oxford: Oxford University Press.
- Schechtman, M. (1996). *The Constitution of Selves*. Ithaca: Cornell University Press.
- Shoemaker, S. (1984). Personal identity: A materialist’s account. In S. Shoemaker and S. Swinburne (Eds.), *Personal Identity: Great Debates in Philosophy* (pp. 67–132). Oxford: Blackwell.
- Sidgwick, H. (1907). *The Methods of Ethics*. London: Macmillan.
- Strickland, L. H., Lewicki, R. J., and Katz, A. M. (1966). Temporal orientation and perceived control as determinants of risk-taking. *Journal of Experimental Social Psychology*, 2(2), 143–51.
- Suhler, C., and Callender, C. (2012). Thank goodness that argument is over: Explaining the temporal value asymmetry. *Philosophers’ Imprint*, 12(15), 1–16.
- Sullivan, M. (2018). *A Theory of Rational Planning and Personal Persistence*. New York: Oxford University Press.
- Tarsney, C. (2017). Thank goodness that’s Newcomb: The practical relevance of the temporal value asymmetry. *Analysis*, 77(4), 750–9.
- Trope, Y., and Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440–63.
- Van Boven, L., and Ashworth, L. (2007). Looking forward, looking back: Anticipation is more evocative than retrospection. *Journal of Experimental Psychology*, 136(2), 289–300.

- Van Boven, L., Kane, J., and McGraw, A. P. (2010). Temporally asymmetric constraints on mental simulation: Retrospection is more constrained than prospection. In K. D. Markman, W. M. P. Klein, and J. A. Suhr (Eds.), *Handbook of Imagination and Mental Simulation* (pp. 131–47). New York: Psychology Press.
- Van Boven, L., and Caruso, E. M. (2015). The tripartite foundations of temporal psychological distance: Metaphors, ecology, and teleology. *Social and Personality Psychology Compass*, 9(11), 593–605.
- Williams, B. (1970). The self and the future. *Philosophical Review*, 79(2), 161–80.
- Zimmerman, D. W. (2007). The privileged present: Defending an ‘A-theory’ of time. In T. Sider, J. Hawthorne, and D. Zimmerman (Eds.), *Contemporary Debates in Metaphysics* (pp. 211–25). Malden: Wiley-Blackwell.