

## **Temporal Fictionalism for a Timeless World**

*Sam Baron, Kristie Miller & Jonathan Tallant*

*Philosophy and Phenomenological Research 102 (2):281-301 (2019)*

### **Abstract**

Current debate in the metaphysics of time ordinarily assumes that we should be realists about time. Recently, however, a number of physicists and philosophers of physics have proposed that time will play no role in a completed theory of quantum gravity. This paper defends fictionalism about temporal thought, on the supposition that our world *is* timeless. We argue that, in the face of timeless physical theories, realism about temporal thought is unsustainable: some kind of anti-realism must be adopted. We go on to provide an argument against eliminativism about temporal thought. While it doesn't follow from this argument that fictionalism about temporal thought is true, we suggest that this nonetheless shows that fictionalism should be regarded as the preferred view.

### **1. Introduction**

Our aim in this paper is to motivate a version of fictionalism about temporal thought. Temporal thoughts are everyday thoughts such as Penny's thought that the baby shower occurred later than her birthday; Barry's thought that it will rain tomorrow; Annie's thought that the future looks grim; Freddie's thought that it took 5 seconds for her to come up with a counterexample; Helen's thought that it is not possible to change the past; and so on. These temporal thoughts seem to make indispensable reference to time. If, however, our world is timeless, then how are we to make sense of such thoughts?

We should be clear from the outset that in this paper we assume that there are temporal thoughts even if time does not exist. For this is how our world is: it is full of temporal thoughts. And so, if a timeless theory is true, then we live in a world that is both timeless and full of such thoughts. In short, the assumption that a timeless theory is (or at least could be, for all we know) true requires assuming (at a minimum) that the theory is compatible with conscious experience.

Thus, our claim in this paper is a conditional one: if certain physical theories, according to which there is no time, are true, then temporal fictionalism should be a leading view about our temporal thought. In (§2) we say more about why one might think the antecedent of this

conditional should be countenanced, but our primary aim is to argue for the consequent. Before we do so, however, we should say something about the scope of this paper.

We cannot hope to show that temporal fictionalism is the only, or even the best, view, if the relevant physical theories turn out to be true. If those theories are true, then things are very unlike we otherwise suppose. For instance, it may well be that if such theories are true, then our world lacks causation. We consider this possibility in §5.1. If that is so then that, in turn, makes it difficult to see how to make sense of many ordinary notions up to and including our being persons. So, in this paper we need to come up with some ways to make sense of some of these ordinary notions in the service of elucidating and motivating fictionalism. The task of fully explicating these notions and defending that they are the best ones on offer is part of a much larger project. What we do here is begin the project of figuring out how to accommodate temporal thought if the relevant physical theories are true.

It is natural to see the question of how to make sense of temporal thoughts if our world is timeless, as broadly analogous to one that arises for certain meta-ethical positions that deny the existence of moral facts. Supposing that there are no mind-independent moral facts, how are we to understand everyday moral thought? The last century saw a proliferation of answers to this question (error theory, fictionalism, realism, non-cognitivism, quasi-realism and so on) and to a further debate over which of these might be the correct approach to ordinary moral discourse. Instances of moral error theory must make sense of the seeming efficacy and ubiquity of moral thought. Any putative instance of *temporal* error theory will presumably have to do something very similar.

Rather than re-inventing the wheel, we will start with a space of options for making sense of temporal thought in the absence of time that is similar to the space available for making sense of moral talk without moral facts. We begin by explaining why temporal error theory is a view we should countenance, (§2), before moving on to show that realism and non-cognitivism about temporal thought appear difficult to sustain (§3-4). We go on to consider the contrast between two important anti-realist approaches—fictionalism and eliminativism—and present two arguments against eliminativism (§5).

## **2. Temporal Error Theory**

Temporal error theory is a response to the radical claim that time does not exist. The status of time has been called into question in recent work on quantum gravity. According to Huggett and Wüthrich (2013), a number of approaches to quantum gravity appear to deny the

fundamental existence of time. In particular, loop quantum gravity and the causal sets approach both seem to lack a spatiotemporal metric at the fundamental level. Huggett and Wüthrich (2013) also maintain that the same may be true for string theory, given the existence of theory duals of string theories that lack a fundamental spatiotemporal structure.

Despite the apparent loss of time at the fundamental level, a number of physicists and philosophers of physics have urged that time is nonetheless an emergent phenomenon, arising at a non-fundamental level. Wüthrich (forthcoming) differentiates between two kinds of temporal emergence. First, there is emergence of an inter-theoretic kind. Crowther (2016, 2017) argues that this kind of emergence involves a relationship of deducibility between theories. Thus, to say that time is emergent in this sense, is to say (roughly, and in part) that the laws of general relativity can be derived from the laws of a more fundamental theory (such as loop quantum gravity in its dynamical form) via bridge laws. This style of emergence sits entirely at the level of theories and describes the way in which one theory may be seen as arising out of another.

Second, there is a more metaphysical or ontological style of emergence. This notion of emergence is not a relationship between theories. It is, rather, a relationship between aspects of physical reality. To say that time emerges from something more fundamental in this metaphysical sense is to say that time depends for its existence on some more fundamental ontology; presumably whatever ontology is posited by a theory of quantum gravity (such as whatever physical objects are represented by spin-networks, or by causal sets).

There is a rising consensus that time is emergent in at least one of these two senses.<sup>1</sup> However, a concern is raised by Maudlin (2007). Maudlin argues that in order to recover time, it is not enough to show that time, or some mathematical structure isomorphic to spacetime, can be derived from a more fundamental theory.<sup>2</sup> In addition to the derivation, we need some reason to suppose that the derivation is *physically salient*.

---

<sup>1</sup> There is, however, at least some room for doubt in both cases. With respect to the first sense of emergence, it remains unclear whether time can be derived from a more fundamental theory that is timeless. For instance, Wüthrich (forthcoming) (following Rovelli (2004, 2011)) argues that one can derive something like space from kinematical loop quantum gravity, but concedes that recovering time may be more challenging. The dynamical form of loop quantum gravity -- the spin foam model -- is not as well understood, and so it is less clear how the derivation might proceed (see also the discussion of loop quantum gravity in Lam and Wüthrich (2018)). There is also a lingering worry that any derivation of time that is successful, is only successful because time is tacitly being assumed in one form or another. It may be that one only gets time out if, in at least some sense, one puts time into the relevant derivation. It remains unclear whether one can achieve a fully successful derivation of time from 'no time'.

<sup>2</sup> Maudlin's criticism was originally raised in the context of wave function realist accounts, which seem to also deny the fundamental existence of spacetime. As Huggett and Wüthrich (2013) note, however, the objection can be extended to timeless approaches to quantum gravity (and is all the more pressing for being generalised in this fashion).

It is somewhat unclear what Maudlin means by 'physical salience'. We can, however, take Maudlin's point to concern the second notion of emergence outlined above. In order for a derivation of time to be physically meaningful, we need some account of how time as a physical phenomenon might depend for its existence on some more fundamental, non-temporal reality. The trouble with appealing to mere derivation is that it leaves the physical-cum-metaphysical picture of time's emergence opaque. The worry, then, is that derivation alone is not enough to establish the proper emergence of time. The inter-theoretic relation of emergence must be accompanied by something with a more metaphysical flavour.<sup>3</sup>

The difficulty is that it is unclear how to characterise the metaphysical relationship between time and an underlying timeless reality. Standard conceptions of metaphysical emergence, familiar from the philosophy of mind and the philosophy of science, make use of metaphysical notions of *composition* or *constitution*. As Ney (2015), Lam and Wüthrich (2018) and Wüthrich (2017) have all argued, however, it is unclear how composition or constitution might be applied to the relationship between a higher-level temporal reality and an underlying timeless reality.<sup>4</sup> Of course, there are relations in metaphysics other than composition that one might use to account for the emergence of time. According to Lam and Esfeld (2013), however, a number of these also fail. In particular, they argue that causal, temporal and supervenience relations of all kinds won't work as a basis for characterising the metaphysical relationship between time and no time.

This, in turn, has led some philosophers to appeal to a functionalist account of temporal emergence. The idea, roughly, is to analyse spacetime in functional terms, and then locate objects in the underlying non-spatiotemporal ontology that can play the relevant spacetime role. So, for instance, based on Knox's (2013, forthcoming) work, Lam and Wüthrich (2018) argue that spin-networks in loop quantum gravity can play a localising role associated with

---

<sup>3</sup> Huggett and Wüthrich (2013) respond to Maudlin (2007) by arguing that a notion of physical significance can be defined for timeless theories, so long as we let the theory itself set the standard for what counts as physically significant (rather than using a pre-theoretic criterion for physical significance that is external to the theory). While we cannot argue the point here, we believe that reading Maudlin's objection in a more metaphysical spirit as we have done here avoids the Huggett and Wüthrich (2013) style response. Even if it doesn't, Wüthrich (forthcoming) at least appears to agree with the importance of the metaphysical dimension of emergence for developing an account of temporal emergence.

<sup>4</sup> One difficulty concerns individuation. If we use composition, say, to characterise the relationship between time and an underlying reality that is not even spatiotemporal, then we need to say that time itself, or temporal phenomena more generally, are composed of parts that lack a spatiotemporal location. Without a spatiotemporal location, however, it is unclear how to individuate parts, or how to talk about the way in which parts might be arranged into some whole. Individuation, for parts, is at least partly a matter of their spatiotemporal location, as is their arrangement. Though see Le Bihan (2016, 2018) for an attempt to revive a composition-based approach to timelessness. Note, however, that Le Bihan's approach is not supposed to be emergentist in nature (see Le Bihan (forthcoming)).

spacetime (the role, roughly, of spatiotemporally locating some object). They argue for this claim on the basis that the volume and area operators over spin-networks agree with the metric expression of general relativity at low energies (the metric expression of the relevant operators diverges from general relativity diverges in the high energy regime).<sup>5</sup>

Whether the spacetime functionalist approach ultimately works, remains to be seen. Ney (2015) has criticised a view along these lines on the grounds that it fails to provide a picture of how spatiotemporal phenomena are constituted by something more fundamental. The worry, we suppose, is that spacetime functionalism might only yield the appearance of there being time: it shows how there can be something that can do the work that time does. This doesn't necessarily show that there is time at some higher level. For that, one would need to show how time is made up of something more fundamental.

We are not sure whether Ney is right about this. On at least some versions of functionalism, playing the functional role of time is sufficient to say that time exists. Still, we do have some lingering doubts about whether functionalism yields a genuine picture of emergence of something temporal from something that is non-temporal. If time just is whatever plays the time role, then that would seem to suggest that the underlying ontology of a timeless theory has temporal properties after all. It remains unclear whether what is captured is a genuinely stratified picture with time above and no time below.

To this we might add that there is something conceptually difficult about the idea that time emerges, and that this conceptual difficulty may underwrite some of the worries that philosophers like Maudlin and Ney have about emergent time. The worry is similar to concerns about the emergence of consciousness. Just as it can be difficult to see how anything with subjective qualities can arise out of the purely non-subjective nature of the physical world, equally it is difficult to see how time can come from no-time. In both cases, the underlying metaphysical basis seems so different in kind to the emergent phenomenon, that the very idea of emergence appears vertiginous. Of course, in both cases, this may speak more to our epistemic limitations, or even (more specifically) to the present authors' failure of imagination. Still, we hope that the idea that time metaphysically arises out of something completely non-temporal gives the reader at least some pause, and makes them wonder how this could be so.

The apparent conceptual gulf between time and a more fundamental timeless reality is, we believe, enough to motivate the consideration of alternative views. In particular, it is worth considering whether we cannot, after all, simply deny the existence of time outright, and thus

---

<sup>5</sup> This assumes a particular approach to the classical limit of the theory; see Oriti (2014) for an alternative.

sidestep the messy business of providing a metaphysical account of temporal emergence (for one attempt to develop a picture along these lines, see Le Bihan (forthcoming)). Consider the analogy with consciousness. In light of the apparent conceptual gap between the physical and the mental, some have considered whether we can do without phenomenal consciousness completely. Such a view might seem a bit crazy to most philosophers, but it is worth exploring (and has been explored in detail). It is roughly in this vein that we will proceed in this paper. We explore what to say about a case where some physical theory that denies the fundamental existence of time is true and, furthermore, that time is not an emergent phenomenon either. The resulting picture of reality is what we will call *deeply timeless*.<sup>6</sup>

Against the backdrop of deep timelessness, we can then consider our ordinary temporal thoughts. Since these thoughts seem to make indispensable reference to time, how are we to make sense of them if time is unreal? Recall that we are assuming that there are temporal thoughts even if time does not exist, since our world is full of such thoughts. We recognise, of course, that the existence of thought itself seems to depend on the existence of various temporal processes, from the most general (like causal relations between mental states at different times) to the most specific (such as rates of neural firing in particular regions of the brain). Without time, it is unclear how the physical basis for these thoughts can exist. It is for this reason that any deeply timeless theory will need to be coupled with some kind of reconstruction project in which it is shown how our experiences, thoughts, observations, and observable entities, can all be accommodated in a timeless setting (Healey 2002). While these recovery projects are as yet incomplete, they typically proceed by first recovering our phenomenal experiences (see, for example, Barbour 1997). That is, they try to show that our phenomenal experiences supervene on some aspect of the structure of the world that does not involve time. They then attempt to

---

<sup>6</sup> It may be that any of the timeless theories discussed above will be of this kind if a satisfactory account of the metaphysics of how time emerges cannot be given. In addition, Barbour's (1994a, 1994b, 1997, 1999) Machian approach to canonical quantum gravity appears to eliminate time, without any attempt on Barbour's part to show that time emerges. Looking beyond physics, a deep approach to timelessness can also be found in Kant's work. For Kant, time is a conceptual category that humans use to construct the world of appearances. Beyond this conceptual role, however, time has no independent reality. As Kant puts the point in the Critique:

If we abstract from our mode of inwardly intuiting ourselves---the mode of intuition in terms of which we likewise take up into our faculty of representation all outer intuitions---and so take objects as they may be in themselves, then time is nothing. It has objective validity only in respect of appearances, these being things which we take as objects of our senses. It is no longer objective, if we abstract from the sensibility of our intuition, that is, from that mode of representation which is peculiar to us, and speak of things in general. Time is therefore a purely subjective condition of our (human) intuition (which is always sensible, that is, so far as we are affected by objects), and in itself, apart from the subject, is nothing. (A35/B52)

In a similar vein, Healey (2002) argues in favour of a 'secondary qualities' approach to time as a response to the apparent timelessness of fundamental physics. Healey's suggestion is to treat time as a secondary quality in the Lockean sense: as something akin to a response-dependent phenomenon.

bootstrap from the recovery of these experiences, to the recovery of contentful thought, again without appealing to time. There are questions about how successful these parts of the reconstruction actually are.<sup>7</sup> However, in what follows we suppose that this last step of the recovery succeeds. We are interested in how to understand temporal thought *if* a deeply timeless physical theory is true and so we are interested in what to say about temporal thoughts in a deeply timeless world, given that there are such thoughts.<sup>8</sup>

### 3. Introducing Temporal Realism and Anti-realism

We are interested in domains of thought. In this case, the domain of *temporal* thought. That domain is one about which one can be a realist or anti-realist. We will suppose that a domain of thought is *realist* just in case the thoughts in that domain are truth-apt and the core

---

<sup>7</sup> Not everyone thinks that such states could exist in a timeless world. Braddon-Mitchell and Miller (forthcoming) argue that contentful states require causal connections, and that the only accounts of causation in timeless worlds (see for instance Baron and Miller 2014) are ones that tacitly appeal to contentful mental states. Hence any timeless theorist is caught in a vicious circle: to vindicate causation she needs to appeal to contentful mental states, but to vindicate contentful mental states she needs to appeal to causation. That's fine. Our project is focused, not on explaining how there could be such thoughts in a timeless world, but in figuring out what to say about such thoughts if they do exist. We concede that it's a lacuna of our current argument that we don't reply directly to Braddon-Mitchell and Miller.

<sup>8</sup> We recognise, also, that deeply timeless theories face a serious threat of empirical incoherence. Following Barrett (1999), we suppose that a theory is empirically incoherent when the truth of such a theory undermines the empirical justification for such a theory. A deeply timeless theory faces a threat of empirical incoherence because it appears to deny the existence of any spatiotemporally located entities. An observation that confirms a theory, however, is always an observation of some such entity. Accordingly, if such an observation were to be made in favour of a deeply timeless theory, then it would follow that the theory is false, since the observation itself would seem to require positing physical phenomena that the theory denies exist. The truth of the theory thus precludes its confirmation. Huggett and Wuthrich (2013, p. 276) put the point in terms of local beables (roughly, spatiotemporally located entities):

Thus, goes the worry, if a theory rejects the fundamental existence of spacetime, it is threatened with empirical incoherence because it entails that there are, fundamentally, no local beables situated in spacetime; but since any observations are of local beables, does not it then follow that none of our supposed observations are anything of the kind?

If time is emergent then the problem of empirical incoherence can be addressed straightforwardly. The emergence of time brings with it the potential for spatiotemporally located entities to exist, and from there empirical confirmation can proceed in the usual way. If time does not emerge, however, then it is unclear how the problem might be addressed. Perhaps it can't. At any rate, we admit that we have no solution to the problem. It doesn't follow from this, however, that a deeply timeless theory cannot be true. All that follows is that such a theory cannot be confirmed via observation. It doesn't even follow that a deeply timeless theory cannot be justified. For there may be non-empirical sources of justification of the theory, of the kind that we use in philosophy (and, indeed, elsewhere in science, for empirically equivalent theories). And so we must be careful to separate the truth of a deeply timeless theory from its empirical confirmation. As noted, to be true, such a theory should (minimally) be compatible with the existence of conscious experience. To be confirmable, the theory should provide a basis for empirical observation. We will set confirmation aside and focus only on truth. It is with respect to the potential truth of such a theory, then, that we raise the question of what to say about temporal thoughts.

thoughts<sup>9</sup> in that domain are true.<sup>10</sup> A domain of thought is error theoretic just in case the thoughts in that domain are truth apt and the core thoughts in the domain are false. A domain of thought is non-cognitive just in case the core thoughts in that domain are not truth-apt,<sup>11</sup> because those thoughts are not beliefs. Thus, we can understand realism, error theory, and non-cognitivism about the domain of temporal thought, respectively, as follows:

**Temporal Realism:** Ordinary temporal thought is truth-apt, and many core thoughts in that domain are true.

**Temporal Error Theory:** Ordinary temporal thought is truth-apt and core thoughts in that domain are not true.

**Temporal Non-Cognitivism:** Core thoughts in the temporal domain are not truth-apt because such thoughts are not cognitive.

In what follows, we will set aside non-cognitivism. While we believe non-cognitivism to be an interesting option to pursue, in the temporal case it is difficult to see what our temporal thoughts might be if not beliefs.<sup>12</sup> In particular, there appears to be a paucity of attitudes that might form the basis of a mature temporal expressivism.<sup>13</sup> This leaves us with realism and error

---

<sup>9</sup> Even moral error theorists want to allow that some moral thoughts are true (for instance, they want to allow that ‘it’s not the case that lying is wrong’ is true) and want to allow that thoughts that are part of moral thinking (such as thoughts about what causes suffering etc.) are true. Joyce, for instance, draws the distinction between atomic and non-atomic moral propositions, and holds that moral error theory is the view that all atomic moral propositions are false. We are attempting to draw the same kind of distinction here, but without appealing to purely syntactic features (such as atomicity). Roughly, one can think of core moral thoughts as those about moral properties or facts (and which such facts obtain). Similarly, core temporal thoughts are, again roughly, those about temporal properties or facts. Our aim here isn’t to rigorously define the difference between core and non-core thoughts, but simply to pay heed to there being some such difference, and note that whatever the difference amounts to, it can be transposed over to the temporal case as well.

<sup>10</sup> This allows that some thoughts in the domain can be false, but that their being false is not sufficient to render the domain non-realist.

<sup>11</sup> This allows that some thoughts in a domain might be truth-apt, and yet the domain as a whole is still non-cognitivist.

<sup>12</sup> Deng (2018) links the temporal and normative cases together quite closely. The account of temporal ontology that Deng offers may provide the resources needed to develop a version of non-cognitivism.

<sup>13</sup> Since, arguably, the most well accepted and sophisticated versions of moral non-cognitivism on the market are variants of expressivism, it seems natural to develop temporal non-cognitivism in its guise of temporal expressivism. The question becomes *which* non-cognitive states the expressivist might identify with our temporal thoughts. There are roughly two kinds of options: emotive attitudes such as regret, anticipation, nostalgia, fear, anxiety, and evaluative attitudes such as desire or preference, all of which display ‘temporal’ asymmetry (Callender and Suhler 2012). Suffice it to say that whichever attitudes one chooses—emotive or evaluative—one significant problem is that there do not appear to be sufficiently many, or sufficiently fine-grained, non-cognitive attitudes to capture all of the different temporal thoughts we have. To illustrate: we can think ‘it was raining an



theory. Temporal realism is viable if the part of the world that timeless theories eliminate is not what makes ordinary temporal thought true. As we see it, there are two ways to satisfy the antecedent of that conditional. First, it could be that what physicists mean by ‘time’, is not what ordinary folk mean; what is eliminated in a timeless physical theory is not what ordinary temporal thought is about. Perhaps *folk time*, as we might call it, exists, even if *physical time* (the thing timeless physical theories eliminate) does not (one might think that something like this is the right story about solidity: physics told us that there is no physical solidity, not that there is no folk solidity. The case of colour is another example. Arguably, colour science tells us that there are no colours *in the folk sense of colour*). Call a version of Temporal Realism that embraces this view, *Folk Temporal Realism*. On this view it is folk time, and folk temporal relations, that make true ordinary temporal thought, even though the time of physics has been eliminated.

What *is* folk time? It’s not clear what content we should attribute to the folk concept of time, so it’s not clear whether timeless physical theories eliminate folk time. The best we can do, borrowing a strategy from Baron and Miller (2015) and Tallant (forthcoming-b), is consider a range of metaphysical conceptions of time, and see what, if anything, these conceptions share. Since metaphysical conceptions of time are generally taken to be somewhat sensitive to folk intuitions about temporality (in at least the sense that they capture some pre-theoretical insight into the nature of time), it is reasonable to suppose that whatever these views share, this will be a necessary (though likely not sufficient) condition for there being folk time. If, whatever that condition is, it is absent in a timeless world, then we can conclude that such worlds contain neither physical *nor* folk time.

In the metaphysics of time there are, broadly, three such pictures of temporal reality: the A-theory, the B-theory and the C-theory. Very roughly, the distinction between these three views is as follows. According to the B-theory, time is to be understood via the B-series: an ordering of events in terms of the relations of earlier-than, simultaneous-with and later-than. The B-theorist maintains that there is nothing more to time than an ordering of this kind.

According to the A-theory, by contrast, time is to be understood via the A-series: an ordering of events in terms of whether those events are objectively past, present or future, and where there is a persistent, global change in which events possess these objective

---

hour ago’, ‘it was raining two hours ago’, ‘it was raining three hours ago’, and so on. Since the object of these thoughts is the same in each case—that it rains—what distinguishes them must be the conative attitude we have towards that state of affairs. There aren’t sufficient attitudes to allow us to distinguish the three thoughts listed, let alone all of the thoughts we can have about its having been raining at past times.

determinations. The A-theorist therefore maintains that the B-series does not exhaust the nature of temporal reality. The A-series is required to give a complete account of time.

More recently, a third picture of temporal reality has emerged according to which time is characterised by a C-series.<sup>14</sup> Where the B-series is constituted by relations of earlier-than and later-than which confer a direction to time (from earlier, to later) the C-series is a non-directed ordering of events in terms of C-relations, (which, for present purposes we can think of as relations of ‘betweenness’ that obtain between triples of events so that one event, E is between E\* and E\*\*). All three series generate an ordering of events, but the C-series is a non-directed ordering, the B-series is a directed ordering, and the A-series is a dynamical directed ordering in which there is a change in which moment is present.

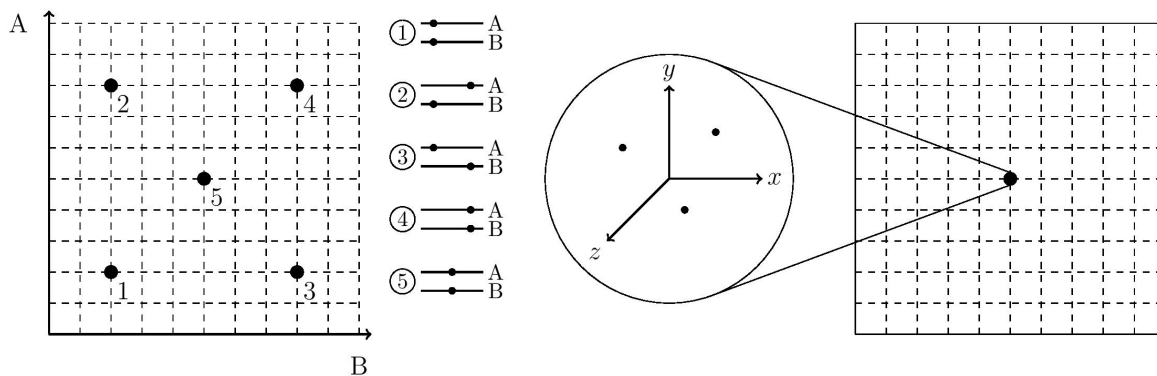
For our purposes, what matters is that all three conceptions of time assume (minimally) that an objective temporal ordering – be it the A-series, the B-series or the C-series – is necessary for the existence of time. In what follows, we will suppose that there being a C-series is necessary (but perhaps not sufficient) for the folk concept of time to be satisfied. Hence, if timeless physical theories dispense with a C-series, they also dispense with folk time. As it turns out, this is precisely what a number of timeless physical theories dispense with (see Huggett and Wuthrich's (2013) general discussion, as well as Barbour's (1994a, 1994b, 1997, 1999) Machian approach to quantum gravity and Rovelli's (2007, 2009) loop quantum gravity approach). Why so? So far, we've said relatively little about the picture of the world painted by the theories of quantum gravity in which we are interested. In what follows we try to paint this picture a little, to make clear that these theories do jettison even a C-series.

According to some timeless theories of reality (Barbour (1999) is an exemplar), we can best imagine reality as a consisting of a space of large three-dimensional objects—every nomologically possible way our world can be, at a time. These are sometimes known as instants, since they have no temporal extension, but it is better to think of them simply as three-dimensional configurations in a configuration space, where each configuration is a specification of the location of every particle and the inter-particle distances between them. Importantly, according to these timeless theories although all of the events/objects on a particular instant bear spatial relations to one another, events/objects on different instants bear no temporal relations to one another. That is to say, there is no ‘right way’ to order these instants via temporal relations; nor is there any sense in which we can measure temporal distance

---

<sup>14</sup> Of course, McTaggart introduced the C-series more than 100 years ago: but he didn't think of this as a temporal ordering in the way that contemporary C-theorists do.

between events/objects located at different instants. So not only is it not the case that there is a temporal ordering of events from earlier, to later that is directed, there isn't even an undirected temporal ordering of events: there is no C-series.<sup>15</sup>



**Figure 1:** Configuration Spaces. The left-hand graph depicts a configuration space for two particles – A and B – that are situated along a line in a one-dimensional space. The numbered nodes each represent a different configuration for the two particles with respect to how they are situated along the line. The right-hand graph represents a  $3N$  dimensional configuration space. Each point is a 3D spatial configuration of  $n$  particles. Points in the space have spatial dimensions, but there is no metric structure over them. If configuration space is the fundamental arena of reality, rather than spacetime, then there are many 3D configurations, and no temporal distance metric. The points can be ordered according to some metric or other, but in the configuration space posited by timeless theories there is no spatiotemporal metric structure that can be used to yield a C-series.

If the ‘arena’ of reality is a configuration space of this kind, then it would appear that Folk Temporal Realism is in trouble. Such timeless physical theories dispense with a C-series, and hence dispense with folk time and folk temporal relations. There is only a configuration space without any spatiotemporal ordering. Thus, given that Folk Temporal Realism is committed to the view what makes temporal thought true is (folk) time, and (folk) temporal relations, it follows that Folk Temporal realism is false: none of our core temporal thoughts are true.

Of course, there may be other options for developing realism. To the best of our knowledge, there is no detailed study of ‘the folk’ with regard to their conception of time. If such data can be gathered, that may change the picture here. Nonetheless, drawing on mainstream metaphysical models and treating these as at least partially representative of a folk-conception of time, Folk Temporal Realism does not seem to fare well. In the absence of a positive proposal that does better, we will not consider temporal realism any further in this paper.

<sup>15</sup> That being said, Barbour does show how to produce directed orderings from configuration space, some of which may be isomorphic to temporal orderings. Barbour is adamant, however, that while one can define these orderings, they are not physically meaningful. The world is not, in any meaningful sense, actually ordered by these orderings. They are, rather, orderings that one can project onto configuration space.

#### 4. Temporal Error Theory

This brings us back to temporal error theory. Temporal Error Theory divides into two further positions: fictionalism and eliminativism.

**Temporal Eliminativism:** (i) Ordinary temporal thought is truth-apt; (ii) core thoughts in the temporal domain are not true, and (iii) ordinary temporal thought ought to be eliminated.

**Temporal Fictionalism:** (i) Ordinary temporal thought is truth-apt; (ii) core thoughts in the temporal domain are not true, but nevertheless, (iii) ordinary temporal thought constitutes a useful fiction, and one that we should continue to engage with.

We take it that, at least in outline, temporal eliminativism is straightforward enough: it simply calls for the elimination of temporal thought. But for us to make sense of temporal fictionalism, we must first make sense of what a temporal fiction might be.

A temporal fiction is any story about the world according to which propositions such as ‘it took 5 seconds for her to come up with a counterexample’ are true in the fiction, but false otherwise. Temporal fictionalism is thus the view that our temporal thoughts express propositions that are strictly speaking false, but are true in some such temporal fiction. The easiest way to understand a temporal fiction is to imagine a complete description of a world in which time exists—the story that most of us take to be true—and now treat that description as the story of time. We can then talk about what's true in the temporal fiction simply by appending a fiction operator to the propositional content of temporal beliefs in the usual way.

We are not, however, principally interested in this kind of ‘operator’ fictionalism. That’s because we are concerned with how temporal thoughts connect to one another, and to deliberation and agency. So, we are interested in how people *behave* when they token such thoughts, not merely in what sort of operator we could append to the propositional content of the thoughts to make them come out as true.<sup>16</sup> So, the fictionalism we have in mind is *pretence* fictionalism. According to a pretence fictionalism, engaging in a fiction means adopting an attitude of make-believe or pretence toward a particular belief or particular package of beliefs.

---

<sup>16</sup> For more on what we call ‘operator’ and ‘pretence’ fictionalism, and why in general to prefer the pretence iteration of the view, see Armour-Garb and Woodridge (2015: especially pp. 31-6).

That is, one pretends *as if* the relevant belief(s) are true. Which package of beliefs? The beliefs whose contents are such that, when inside the scope of the relevant ‘in the story of time’ operator, those beliefs come out as true, and are false otherwise. According to pretence temporal fictionalism, one engages in a temporal fiction if one pretends that one’s temporal thoughts are true. So, suppose Sara behaves excitedly today: she makes an utterance such as ‘tomorrow is my birthday’ and she makes arrangements for a party, and so on. Suppose, moreover, she believes that time does not exist and thus that, strictly speaking, there is no tomorrow. Thus, she takes ‘tomorrow is my birthday’ to be strictly speaking false. In this situation, Sara is engaging in the temporal fiction. She *pretends* as if her temporal thoughts are true.

Given the choice between eliminativism and fictionalism, what is the best response to the actual world being timeless? In attempting to answer this question, we will take our cue from the moral case. It is common for arguments in favour of moral fictionalism to come in the form of arguments against eliminativism. With such an argument in hand then, given that moral thought is false—as the error theorist supposes—and given that we shouldn’t eliminate it, we have at least some reason to adopt a fictional attitude toward our moral thoughts. Of course, it doesn’t follow from the failure of eliminativism that fictionalism is true, but it does quickly become the best option.

Our arguments against temporal eliminativism concern *agency*, temporal thought, and agentive thought. Temporal thought is tightly connected to a range of other notions. In particular, temporal thought is central to the agential web: a web of beliefs that are necessary for agency, by being necessary for deliberation. These beliefs include a range of beliefs about the past, present and future, as well as beliefs about causation, including what causally precedes what, and beliefs about persons, including beliefs about whether a person remains the same person over time. Underpinning the various concepts that make up the agential web and the thoughts that are partially constituted by those concepts, are our ordinary temporal thoughts and the temporal concepts that partially constitute those thoughts. If ordinary temporal thought is not in good standing it is difficult to see how these other *agentive thoughts* could be in good standing either.

Without wishing to belabour the point, recall that our folk notion is taken to require some notion of ordering—i.e. at the very least a C-series. If we deny that there is any kind of ordering, agency is hard to understand. At a minimum, and setting aside theoretical niceties as to the best theory of action, we take it that human agency will require planning and the

subsequent execution of those plans. The key, of course, is ‘subsequent’. If our metaphysical model lacks subsequence, then it lacks order; if it lacks order, we lack agency.

Here, then, is our first key claim: if error theory about temporal thought is true, then it follows that error theory about agentive thought is also true. For if our temporal thoughts fail to be true, then our agentive thoughts fail to be true, since the truth of the latter rely on the truth of the former.

Our second key claim is this: if we respond to error theory about temporal thought by eliminating that thought, then we must also eliminate agentive thought. For not only does the truth of the latter depend on the truth of the former, but, in addition, the capacity to have thoughts of the latter kind depends on the ability to have thoughts of the former kind. If we eliminate temporal thought we cannot make sense of planning for the *future* and thus of practical reasoning; we cannot make sense of moral accountability for *past* or *future* acts that are ours and others’ actions—and thus the heart is ripped out of moral reasoning. An eliminativist approach to our ordinary temporal thought threatens to force an eliminative attitude to thoughts and concepts making up the agential web.

Assuming that the world is deeply timeless, we then face a choice: eliminate agentive thought, or fictionalise it. We argue in favour of fictionalising it. Fictionalising agentive thought, however, requires also fictionalising the embedded temporal thoughts. Hence our overarching argument is as follows:

- (1) If our world is deeply timeless, then we have reason to adopt fictionalism about agentive thought over eliminativism about agentive thought.
- (2) Fictionalism about agentive thought requires fictionalism about temporal thought.
- (3) Therefore, if our world is deeply timeless, then we have reason to adopt fictionalism about temporal thought over eliminativism about temporal thought.

The challenge in defending the argument lies with premise (1). Why think that timelessness forces one toward fictionalism about agentive thought? Our defence of (1) will take a form that is familiar from work on moral error theory. Usually, an argument against elimination identifies a cost associated with the elimination in question. The cost of elimination is then converted into a benefit of fictionalism. Thus, we should adopt a fictional attitude toward our moral thought, say, because the moral fiction is useful: it allows us to keep our moral thought, which allows us to avoid the cost of removing it (see Joyce 2001). It may be that to engage in such a fiction preserves social cohesion, ensuring the maintenance of a basic social contract. Or we

may have self-interested reasons for everyone to engage in the fiction: after all, if people did not generally engage in the fiction that murder is wrong, then any one person's life would be hostage to the whims of their neighbours. More generally, what it means to say that a fiction is useful is just this: a fiction *F* is useful just when, for prudential reasons, one ought to engage in the fiction, where 'engaging with the fiction' means (at a minimum) that one ought to continue behaving *as if* the fiction were true in the sense advocated by pretence fictionalism.

Matters are more complex in the case under consideration. Error theory about agentive thought follows from error theory about temporal thought. Error theory about agentive thought, however, might be thought to entail error theory about prudential reason. After all, if one's thoughts about what one did do, will do, and might do (and could do) are not true, then one might think that one cannot have prudential reasons to do anything. If one has no prudential reasons, however, then one cannot have a reason to endorse fictionalism over eliminativism of the above kind. We will return to this issue later on, in §5.3. For now, however, we are simply going to assume that one can have prudential reasons even if one's agentive thoughts are systematically false.

## **5. Against Eliminativism**

In order to defend premise (1) in our argument against temporal eliminativism, we must consider two kinds of cases. First, a case in which our world is deeply timeless but nonetheless contains causation. Second, a case in which our world is deeply timeless and lacking in causation. We will argue that in both cases we have good reason to resist the elimination of agentive thought.

### **5.1. Deeply timeless worlds with causation**

Suppose that our world is deeply timeless but contains causation. Perhaps causation is fundamental, and there is no emergent temporal or spatiotemporal structure. Or perhaps causation, but not time, emerges from a fundamentally timeless reality. Exactly how a world might contain causation but not time is not an issue we will consider here. Rather, we will simply assume that there is some sense to be made of the idea.

If time does not exist, then even if causation exists, our temporal thoughts are in error. So, there is still the question of what to say about our temporal thoughts in such a world. Should they be eliminated? As discussed above, temporal thought is necessary for agentive thought

and, moreover, if we eliminate temporal thought we end up eliminating agentive thought. The question, then, is whether eliminating agentive thought is something we can live with.

Perhaps if the gains associated with eliminating both agentive thought and temporal thought were significant enough, then we might have an overall reason to eliminate temporal thought. But we don't see that this is so. What does elimination have going for it? Perhaps if we eliminate temporal thought, then we get some boost to our epistemology. We gain the virtue of bringing our system of beliefs more directly into line with the way the world is. This, in turn, may help us to develop a more accurate scientific understanding of the world, or it may make us feel happier knowing that we are closer to the truth.

This benefit, however, seems to be outweighed by the costs of eliminating agentive thought. One could hardly think that were humans, as a species, to lose the capacity for agentive thought, then this would be no great loss. Many of us will think that this would be tantamount to ceasing to exist if, for instance, one thinks that what it is to be a person or a human just is to have thoughts of this kind (more on this below). Even if one doesn't think this, however, agentive thought plays a central role in our lives. Eliminating agentive thought would require rewriting human self-understanding. This would be a massive undertaking.

The problem is especially acute if causation exists. There are two points to be made: the first is the cost of eliminating temporal thought in this context, the second is the benefit of fictionalising.

The cost: if causation is real in this scenario then we can make a difference to the world: we can cause things to happen. Moreover, things will happen to us, things to which we will very much need to respond. If one can in fact make a difference to the world causally speaking, and the world can make a difference to us, then eliminating agentive thought seems disastrous. If we eliminate agentive thought, then we take away any capacity that we might have to reasonably respond to the environment around us in a meaningful way. Our interaction with the world would be limited to reflex and instinctive impulse, rather than high-level cognition and decision-making. It is hard to overstate how radically this would alter the kinds of lives we would live, and, for many of us, alter them for the worse.

The advantage of fictionalising: our agentive thoughts clearly have a causal dimension to them. Agentive thoughts are as much about causation as they are about time. When we believe that something is an option to be deliberated about, we typically believe that the option is something we can bring about. If we have no causal power over the option, then it is hard to see why it would be the object of any agentive thought whatsoever. And whilst without time, our agentive thoughts won't be wholly true, because the embedded temporal dimension of them



will fail to refer, nonetheless, in the presence of causation, those beliefs will be at least partially true: the causal dimension of them is still viable. So there remains a sense in which the beliefs are *apt* despite not being strictly true: they still track the causal aspect of the world, and tracking the causal aspect of the world is useful for one to achieve certain ends. If they are apt (and so are useful), then eliminating seems the wrong way to go; we are better served by fictionalising.

We anticipate two responses to this line of thought. First, in order for eliminativism about agentive thought to be motivated in the first place, agentive thought must be in error. But, one might argue, if causation exists, then our agentive thought is not systematically in error. Agentive thought, in this situation, is straightforwardly true: it is made true by the fact that causation exists: the idea being that the underlying causal reality scaffolds our agentive thinking. If, however, our agentive thought is straightforwardly true, then there is no threat of elimination brought about by eliminating temporal thought.

The trouble with this response is that if temporal thought is necessary for agentive thought, then a causal theory of agentive thought that renders our agentive thoughts true, also renders our temporal thoughts true. It follows that causation of the kind that can ground agency is *also* capable of grounding time. While we are willing to admit that causation can do this kind of work, the resulting picture is not a timeless one. Since we are interested in a case in which time does not exist, but causation does, it is not clear that realism about agentive thought can be recovered quite so easily. In sum, reintroducing the kind of causation that can make our agentive thoughts true just is to reintroduce time into the picture, albeit in an indirect fashion.

This brings us to the second response to our argument: one might simply deny that temporal thought is necessary for agentive thought. This would free one up to adopt an account of causation that is capable of grounding agentive thought, without fear of reintroducing time into the picture. It is clear, however, that the way in which agentive thought is currently structured makes indispensable use of time. All of our agentive thinking is directed to, or developed around, thoughts of time in some manner. Simply denying that agentive thought is like this won't do.

A more plausible response is to recommend a revisionary approach to agentive thought, whereby all such thoughts are stripped of their connection to temporal thought. Perhaps the role that temporal thought once played in agentive thought can now be played by causal thoughts. Thus, instead of talking about a future option, one might frame agency simply in terms of what one has the power to cause; a choice of options amounts to a decision about how to exercise one's causal powers; things that one can exercise no agency over can be construed as things that one has no causal power to affect.

One point in favour of proceeding in this way is that the proposed revision would retain at least some continuity with our current agentive thinking. The downside of the proposal, however, is that it still requires a sweeping change to a domain of thought that is absolutely central to our everyday lives. We need a pretty good reason to adopt such a change, given just how disruptive and difficult the proposed revision is likely to be. So, in the end, even if causation can make agentive thought true, and even if agentive thought can be severed from temporal thought, it seems far less troublesome to simply retain our temporal thought and leave its connection to agentive thought intact. The potential gains we get from eliminating temporal thought entirely, versus simply pretending as if it is true, don't seem substantial enough to warrant a wholesale revision of thoughts that stem from the agentive web.

## **5.2. No Time and No Causation**

This brings us to the second kind of case: the case in which both causation and time are unreal. In this case, the elimination of agentive thought does not carry quite the same implications as when causation is in the mix. It is one thing to propose eliminating agentive thought when we have causal powers, and when we are caught up in causal networks of various kinds. It is quite another thing to propose eliminating agentive thought when we lack causal powers entirely, and when we are not nestled within a network of causal goings on. In this case the cost of eliminating agentive thought is less obvious, and there may seem to be positive benefits to doing so. After all, if we are wasting energy deliberating about things that we think we can bring about, when in fact we can't bring about anything, then perhaps the right thing to do is reorient ourselves to the new world order. Might we remove our agentive thoughts entirely and, with them, our temporal thoughts?

In fact, the cost associated with eliminating agentive thought is just as high when both causation and time go missing, if not higher. To show this, we will argue for two claims. First, that we have good reason to suppose that something like a desire-first view of personal-identity is the right view when both time and causation are missing. Second, we argue that if a desire-first view of personal-identity is the right view, then eliminating agentive thought entails eliminating the very things we are: persons. Hence it is deeply undesirable for persons to eliminate temporal thought. In schematic form, the argument is as follows:

[1] If, actually, our world is timeless and causationless, then it is necessary for being a person that one holds agentive beliefs.

[2] It is necessary for holding agentic beliefs that one holds temporal beliefs.

[3] Giving up one's temporal beliefs would require giving up one's agentic beliefs which, if the actual world is timeless, would require giving up one's personhood.

[4] So, if actually our world is timeless and causationless, then persons cannot eliminate their temporal beliefs: there is no such thing as an actual person who lacks temporal beliefs.

We have already explained why [2] is true: temporal thought is central to agentic thought. [3] and [4] follow from [1] and [2]. So, the challenge in defending the argument is to defend [1]. Here is our argument for [1]. Once again, in schematic form:

[1a] If actually, our world is timeless and causationless, then the best account of persons is a desire-first account.

[2a] If a desire first account is true, then it is necessary for being a person that one holds agentic beliefs.

[3a] So, if actually our world is timeless and causationless, then it is necessary for being a person that one holds agentic beliefs.

This sub-argument relies on two things: first, that if actually, our world is timeless, then the best account of personal-identity is a desire-first view, and second, that the desire-first view delivers a necessary connection between person-hood and the possession of agentic beliefs. Notice that this argument does *not* require that in order to have persons, one must have agents. For it leaves open that there could be agentic beliefs without there being agents. One might think that for there to be agents is more than just for there to be person-stages – roughly a spatial ‘slice’ of a person, more on this below – that have agentic beliefs. If, for instance, agency requires that the agentic beliefs of one person-stage play a role in determining the agentic beliefs of certain other person-stages, and in determining the actions of those person stages, then having agentic beliefs is necessary, but not sufficient, for agency. Hence neither (2a) nor (3a) above, commit one to thinking that agency is necessary for personhood. That's important to bear in mind since ultimately, even if at the end of this paper you are inclined to

think that there would be persons in our world, were it timeless, and that they would be best seen as engaging in a temporal fiction, you might still resist the claim that these persons are genuinely agents in some full-blooded sense. We won't take a stand on this; we just want to flag that it remains an open possibility.

Part of our goal, then, will be to outline the particular account of personhood implicated in [2a] and explain the necessary connection in question. An important challenge in defending the sub-argument is to defend [1a], which raises the very interesting question of what persons are like in a world without time. In fact, in what follows we won't be able fully to defend [1a]. After all, we can't show that there is no possible account of personal-identity that is not preferable to the one we offer. What we can do, however, is show two things. First, we can show that a desire-first account is compatible with deep timelessness, and second, we can show that other extant theories of personal-identity are not. We think this provides good reason for us to accept a desire-first view of personal-identity conditional on our world being deeply timeless, and so gives us reason to endorse [1a].

Let us take the second point first. All traditional accounts of personal-identity rely on a relation of continuity between bodies, minds, or animals, located at different times. On a four-dimensionalist picture, the relation is used to rope together a range of instantaneous things—persons-at-times—into a four-dimensional object that is a person. In the case of psychological accounts of personal-identity, the continuity relation is a relation of psychological continuity. On this view, roughly speaking, two temporal parts are parts of the same four-dimensional person just in case the mental states of the former are continuous with those of the latter. In the case of bodily accounts of personal-identity, the relation is a relation of physical continuity and in the case of animalist accounts it is a matter of being the same animal.

What all these relations of continuity have in common is causation. It is a necessary condition on two person-stages being psychologically continuous that there is an appropriate causal relation between them. The mental states of one instantaneous slice need to be the causal consequence of the mental states of the other instantaneous slice in order to qualify as stages of the same person. Similarly, in the case of bodily continuity, two instantaneous slices are bodily continuous only if the physical bodies of the two slices are causally connected to one another so that there is a relevant continuity of the same physical structure across the person's lifespan. Animalism is the same: there must be a causal relation in virtue of which two instantaneous slices constitute the same animal, otherwise the slices won't be connected in the way required for the continuity of the animal.

The causal relation is, of course, not sufficient for personal-identity according to any of

these views. But it is necessary. If traditional accounts of personal-identity suppose causation to be necessary for persons, and if deeply timeless worlds lack causation, then according to these views they also lack persons. So, if our world is a deeply timeless world without causation, then according to these traditional accounts it's a world that lacks persons. Since our world clearly does not lack persons, then we must look for another account of personal-identity.

In what follows we argue that a desire-first account is just such an account. It is important to be aware that we are not trying to offer a full-blown defence of an account of personhood that is compatible with deep timelessness. That is another paper in itself. Ours is the more modest aim of arguing that amongst extant theories of personal-identity, a desire-first view is the only one that can accommodate the existence of persons in a deeply timeless and causationless world. Since we are persons, if our world is deeply timeless, then we have reason to think that this is the right account. Of course, this leaves open that some future, as yet unarticulated view of personal-identity will be forthcoming, and that this account will not provide a reason to favour fictionalism over eliminativism.

Let us say that in a timeless world,  $x$  is a person-stage just in case  $x$  is an intrinsic duplicate of a person at a time – an instantaneous slice – in a world with time. As discussed, we are assuming, in this paper, that we can make sense of mental content in a timeless world. So, we can relatively straightforwardly accommodate the idea of there being person-stages in a timeless world, assuming that an intrinsic duplicate of a person stage will be one that shares the same intrinsic physical and mental properties.<sup>17</sup> If ours is a deeply timeless world, then it contains many person-stages.

The desire-first approach to personal identity appeals only to person-stages and to certain relations that obtain between those person-stages, which, *crucially*, in the actual world are the relations around which actual person-stages organise certain of their conative attitudes. What's important about this is that it leaves it entirely open just which relations these are. If actually there are relations of psychological continuity or relations of physical continuity, then it may well be (and likely is) the case that various person-stages organise their conative attitudes around one or other of these relations. But suppose there are no such relations because there are no causal relations. Then unlike standard theories of personal-identity, according to which

---

<sup>17</sup> Of course, you might worry that it cannot share the same mental properties, since the wide content of its states will be different from the stage of which it is an intrinsic duplicate. Remember that we are assuming that the timeless theorist can make sense of mental content *and* that she can make sense of there being person-stages which have the content that our mental states seem to have (after all, it's a hypothesis about how the actual world is). So we will assume that there can be person-stages that have temporal thoughts much like our actual temporal thoughts.

under those conditions there are no persons, desire-first theories will say that actual person-stages must be organising their practical concerns around some other relation (or relations). We will have more to say about what that relation might be shortly. First, we need to say more about the desire-first approach to personal-identity. These views are (unsurprisingly) couched in temporal terms. Thus, we will begin by outlining a desire-first account in the usual terms, before showing how one can make sense of such a view without appealing to temporal machinery.

According to desire-first accounts, what makes a relation the personal-identity relation is not some independent metaphysical (or semantic) matter of fact about personal-identity. Rather than its being the case that some person-stage, P, ought to care about some future person-stage because that stage is a continuer of P, instead, such views hold that what makes it the case (*inter alia*) that the future stage is a continuer of P, is the fact that P has certain attitudes towards that stage, or more precisely, towards the *relation* that the stage bears to P. The idea, roughly, is that there are very many different relations (and ancestrals thereof) that obtain between sets of person-stages, including relations of psychological similarity and causal continuity (so called psychological continuity relations), relations of physical or bodily similarity and causal continuity (so called physical continuity relations) and so on.

Desire-first accounts of personal-identity hold that conative states determine facts about personal-identity.<sup>18</sup> On this view, the conative states of a person-stage determine which of these relations the personal-identity relation is in the following way: the personal-identity relation is the relation around which the person-stage in question *organises* its attitudes of practical care. We will have more to say about what it is to organise one's attitudes of practical care, in moment. First, however, what are these attitudes of practical care? Let's suppose that the conative attitudes that are relevant to determining which relation is the personal-identity relation are of two kinds: desires for *extension* of the self where these extensions bear certain sorts of relations to the current person-stage—these are forward looking conative attitudes—and desires to have been *extended from* (i.e. to have as an *origin*), person-stages that bear certain sorts of relations to the current person-stage—these are backwards looking conative

---

<sup>18</sup> Such views are also sometimes known as conventionalist, and their defenders include, *inter alia*, White (1989); West (1996); Johnson (1989). Braddon-Mitchell and West (2001); Braddon-Mitchell and Miller (2004); Miller (2009; 2013); Robinson (2004). Views of this kind have also been gestured towards in Nozick (1981 p 69), Unger (1990 p66) and Whiting (1986). See also Shoemaker (2003) for a related view according to which “our identity as functioning, well-developed agents is constituted by our nexus of cares.” Eklund (2004) calls views of this kind *self-concern relativism*. Kovacs (2016 and ms) develops a related view that appeals to first-person ‘I’ thoughts rather than conative states.

attitudes. Call the former attitudes of anticipation, and the latter attitudes of origin. Rather than it being the case that person-stages have attitudes of anticipation and origin towards other person-stages because those person-stages bear some particular relation to the person-stage in question, instead, it is the attitudes that are primitive, and some relation, such as psychological continuity, is one that a person-stage cares about *because* the primitive attitudes are ones that track that relation. Where this is so, we will say that the practical cares of the person-stage are *organised around* that relation. Importantly, on this view the property of organising attitudes of practical care is a property had by instantaneous person-stages. A person-stage organises, or structures, its practical care around some particular relation, R, just in case that person-stage desires to be extended from stages that bear the R-relation to it, and desires to be extended by stages that bear the R-relation to it. *Organising*, then, is not a notion that presupposes time: it supposes only that person-stages can have desires of a certain sort.

So, supposing for a moment that time is real, and that a person-stage organises its practical cares around the psychological continuity relation. That person-stage desires to be extended by the relation of psychological continuity, and desires to have come from person-stages connected by psychological continuity. In this situation, a person-stage feels regret, pride, shame, and so on for what past person-stages that are psychologically continuous with it did, and it anticipates, fears, dreads, and so on, the experiences of future person-stages that are future-wise psychologically continuous with it. Relative to that person-stage, the psychological continuity relation is the personal-identity relation. Since other person-stages might have primitive attitudes of anticipation and origin that result in them organising their practical cares around some other relation—say, the relation of physical continuity—then relative to those person-stages that other relation—i.e. physical continuity—will be the personal-identity relation for those person-stages.<sup>19</sup>

On standard versions of the desire-first account, it is assumed that in fact, actual person-stages organise their conative attitudes around a mixture of causal and similarity relations: psychological continuity, or physical continuity or animalistic continuity and so on. That's because it's assumed that there are such relations, and that, as such, the conative attitudes of various person-stages are quite likely organised around some, or all, of these relations.

In principle, however, the desire-first account allows for conative attitudes to be organised around straight similarity relations. One can imagine possible person-stages like this:

---

<sup>19</sup> This leaves it open that even if actually there are causal relations, there might be person-stages that organise their prudential attitudes around relations that are not causal. Indeed, such a view has been defended by Duncan and Miller (2016).

stages who do not organise their conative attitudes around any continuity relation at all. These are person-stages whose patterns of care do not track causal relations, but only psychological similarity, or physical similarity, or some other kind of similarity. Most desire-first theorists don't think that actual person-stages are like this, but they also hold that this is a contingent matter, and thus what the relations are around which person-stages organise their attitudes is an empirical question.

If, however, our world lacks both time and causation then it is very plausible that the relations around which actual person-stages organise their conative attitudes are similarity relations of one kind or another. Why so? Well, recall that earlier we noted that timeless physical theories reject the existence of temporal ordering. Despite giving up on temporal ordering, however, all timeless theories proposed to date are consistent with the existence of similarity relations between physical objects. Thus, while timeless physical theories might not allow for any temporal orderings of physical objects, those theories do at least permit similarity orderings of such objects. Indeed, some timeless theories make use of these similarity orderings as a way of recovering the dynamical laws of the theory (see, for instance, Barbour (1997)). As before, the details of these physical theories are not important for present purposes.

Let's suppose that there are actually such similarity relations if our world is deeply timeless. Call each of these similarity relations an S-relation: then there is the S1-relation, the S2-relation, and so on, for each kind of similarity relation around which actual person-stages organise their practical cares. Then, very roughly, according to the desire-first view, person-stages are parts of the same person just in case those stages organize their practical cares around some Sn-relation, and those stages are Sn-related (or bear the ancestral of being Sn-related).<sup>20</sup> Hence the version of the desire-first account that we have sketched is compatible with a deeply timeless and causationless world. More than this, however, a desire-first account of personal-identity offers the best hope of providing an account of personal-identity in such a world. As noted, attitudes of care can be organised around similarity relations; and if our world is timeless, then plausibly they are organised around such relations. Their being so organised is,

---

<sup>20</sup> In fact matters are more complex than this, since even amongst the stages that are all Sn-related to one another, there might be person-stages who organize their practical cares around something other than the Sn relation. Hence, which stages compose a person will always need to be assessed relative to the attitudes of a *particular* person-stage. Relative to person-stage Penny1, say, all of the stages that are S1-related to Penny1 are stages of the same person. Suppose that Penny2 is S1-related to Penny1. So, by Penny1's lights, Penny2 is a continuer of Penny1. But now suppose that Penny2's practical cares are organised around the S2-relation not the S1-relation. Suppose also that Penny1 is not S2-related to Penny2. Then although by Penny1's lights Penny2 is a *future* stage of her, by Penny2's lights, Penny1 is not a *past* stage of her. In general, given some set of person-stages, there is no one right way of determining which stages compose which person. Rather, which person those stages compose will always be relative to the relation around which each stage in the collection organises its practical cares.



according to desire-first views, sufficient for there being persons. Such an account avoids the need for causation, and uses only resources that are available in deeply timeless worlds.

So far, we have argued that the desire-first account provides the best account of persons in a deeply timeless and causationless context. We are now in a position to elucidate the link between the desire-first account and agentic thought.

Suppose that persons are sets of person-stages connected by some relation, *R*, around which those person-stages organise their conative attitudes. In some rough sense, each such person-stage desires that it ‘comes from’ stages that are *R*-related to it, and that it has *R*-related continuers. The relevant bundle of conative attitudes here is quite rich. It includes attitudes of anticipation towards experiences of continuer stages, and attitudes of pride and regret for behaviours of ‘past’ stages. But these attitudes also include various planning and deliberative attitudes. What makes *R* the relation around which a person-stage organises its conative attitudes will be, in part, that stages that are *R*-related to that stage (and are ‘downstream’ of it) are ones about whose interests that person-stage deliberates. Quite generally, prudential attitudes are a crucial component of the parcel of conative attitudes that determine which relation in the world counts, for some person-stage, as the personal-identity relation. That’s because, what it is to organize one’s conative attitudes around relation *R*, is, *inter alia*, to only anticipate the experiences of person-stages to which one is *R*-related, and, as a result, to have a particular kind of care regarding *which* experiences that person-stage has. This latter kind of care involves desires to bring it about that those person-stages have certain kinds of experiences and not others. That, in turn, involves a range of further attitudes, including deliberation, planning, and so on, as ways of trying to bring it about that those person-stages have certain kinds of experiences. Thus person-stages’ having prudential attitudes—attitudes of care of a certain sort about what sorts of experiences some person-stage has—involve having further attitudes, attitudes of deliberating and planning, and these further attitudes involve having agentic beliefs. So, the rich package of conative attitudes that partially ground personal-identity, according to desire-first views, thus requires that each person-stage has agentic beliefs.

Given the link between agentic beliefs and temporal thoughts, it follows that persons must have temporal thoughts. In order to believe that something is an option, say, which is one important kind of agentic belief, one must treat it as something that might happen *in the future*. If agentic beliefs are essential to us being persons, however, then this means that eliminativism about temporal thought is deeply undesirable. For if a person eliminates their temporal beliefs, then they must also eliminate their agentic beliefs, which means giving up

their personhood. If we are essentially persons—as is often supposed—then requiring that we cease to be persons is requiring that we cease to exist. Now, perhaps a being can choose to cease to be a person; perhaps a being can even sometimes rationally so choose. But an account of temporal thought ought not *require* that persons choose to cease to be persons, which is exactly what eliminativism about agentive thought would demand.

### 5.3. Prudence and Preference

To summarise the discussion in this section so far: either a lack of time is accompanied by a lack of causation or it is not. If it is, then we have little reason to eliminate agentive thought and a good reason not to: we need it to manage our own position within a causal network. If time is not accompanied by causation, then eliminating agentive thought comes at the cost of eliminating persons. Either way, eliminativism about agentive thought has little to recommend it. Since eliminativism about temporal thought would lead to eliminativism about agentive thought, the reasons cited for preserving agentive thought carry over to temporal thought as well. The reasons that speak against elimination speak in favour of fictionalism. If we can keep hold of agentive thought, then we can keep hold of persons and we can keep hold of our standard conception of agency. So, we have reason to pretend as if our temporal thoughts are true, so that we can then engage in the same pretence with respect to our agentive thoughts.

As noted in §5.1, we have assumed throughout that the elimination of agentive thought does not also imply the elimination of prudential reasons. Thus, one can still have a reason to adopt fictionalism. But what if error theory regarding both agentive and temporal thought also leads to the loss of all prudential reasons? We won't have anything to say, here, about why this might be so, but one need only think that prudential reasons are in some way grounded in our collective practices and capacities as agents, to worry that if our agentive thoughts are systematically in error, then there are no prudential reasons. Let's suppose one thinks something along these lines. Still, even if there are no prudential reasons, there are preferences, and we retain our preferences. Everything we have said about the importance to us of agentive thought still holds true when thinking in preferential terms. We *desire* to be the kind of thing that has agentive thoughts. Indeed, this desire to have agentive thoughts might be one of our most fundamental. Given this desire, most of us will in fact prefer fictionalism about both temporal and agentive thought, to eliminativism about such thought. Indeed, all this preference really need amount to is that faced with the belief that temporal error theory is true, we would *in fact* not eliminate that thought, nor agentive thought. Instead, we would continue to act as

though such talk were true: and that's all it is to prefer fictionalism to eliminativism.

## 6. Conclusion

We have suggested that an error theory about temporal thought looks most plausible in the context of the truth of a timeless physical theory. In that context, we have argued against eliminativism. If we cannot, or if we prefer not to, eliminate our temporal thought, then once we know such thought to be false the most obvious response is to change our attitude towards such thought: to take such thought to be constitutive of a fiction in which we are engaged. Hence, we have suggested that temporal fictionalism is the best way forward if a timeless physical theory is true.

There is no doubt that deeply timeless worlds are odd: they are not at all the way we intuitively suppose our world to be. And yet serious physicists take timeless theories seriously, and philosophy should do the same. Still, such theories require that we give appropriate consideration to how to make sense of the sorts of beings we are, conditional on such an account being true. Here, we have focussed on making sense of perhaps the most difficult aspect of our mental lives should a deeply timeless theory prove to be true: our temporal thoughts. We have defended, as the best option, a version of temporal fictionalism. Nothing we say here is decisive. Much remains to be said about how to understand a variety of notions, including personal-identity. Still, we hope to have begun this project, and to have at least motivated temporal fictionalism and said something about why we think this view is preferable to other contenders.

## References

- Anderson, E. (2006) Relational Particle Models: 1. Reconciliation with Standard Classical and Quantum Theory. *Classical and Quantum Gravity*, 23(7), 2469–2490.
- \_\_\_\_\_. (2009) Records Theory. *International Journal of Modern Physics D*, 18(4), 635–667.
- \_\_\_\_\_. (2012a) Problem of Time in Quantum Gravity. *Annalen der Physik*, 524(12), 757–786.
- \_\_\_\_\_. (2012b) The Problem of Time in Quantum Gravity. In *Classical and Quantum Gravity: Theory, Analysis and Applications*. Vincent. R. Frignanni (Ed.), pp. 1–25 (New York: Nova).
- Armour-Garb, B. and Woodridge, J.A. 2015. *Pretense and Pathology: Philosophical Fictionalism and its Applications* (CUP: Cambridge)

- Barbour, J. (1999). *The End of Time*. (Oxford; New York: Oxford University Press).
- \_\_\_\_\_. (1994a) The Timelessness of Quantum Gravity: I. The Evidence from the Classical Theory. *Classical Quantum Gravity*, 11(12), 2853–2873.
- \_\_\_\_\_. (1994b) The Timelessness of Quantum Gravity: II. The Appearance of Dynamics in Static Configurations. *Classical Quantum Gravity*, 11(12), 2875–2897.
- Barbour, J., and Isham, C. (1999) On the Emergence of Time in Quantum Gravity. In *The Arguments of Time*. Jeremy Butterfield (Ed.), 111–168. (Oxford; New York: Oxford University Press).
- Baron S. and K Miller (2015) Causation *Sans* Time. *American Philosophical Quarterly*. 52(1): 27-40.
- \_\_\_\_\_. (2015) What is temporal error theory? *Philosophical Studies*. 172 (9): 2427-2444
- \_\_\_\_\_. (2014) Causation in a timeless world. *Synthese*. Volume 191, Issue 12, pp 2867-2886  
DOI 10.1007/s11229-014-0427-0.
- Barrett, Jeffrey A. (1999), *The Quantum Mechanics of Minds and Worlds*. Oxford University Press, New York.
- Braddon-Mitchell, D and K Miller (forthcoming). Quantum Gravity, Timelessness, and the Contents of Thought. *Philosophical Studies*.
- \_\_\_\_\_. (2004) How to be a Conventional Person. *The Monist* 87(4), 457-474.
- Braddon-Mitchell, D and C West (2001) Temporal Phase Pluralism. *Philosophy and Phenomenological Research* 62(1), 59-83.
- Crowther, Karen (2016), *Effective Spacetime: Understanding Emergence in Effective Field Theory and Quantum Gravity*. Springer, Heidelberg.
- Crowther, Karen (2017), “Inter-theory relations in quantum gravity: correspondence, reduction, and emergence.” *Studies in History and Philosophy of Modern Physics*, 1–12.
- Suhler, C. and C Callender (2012) Thank Goodness that Argument is Over: Explaining the Temporal Value Asymmetry. *Philosophers’ Imprint* 12(15), 1-16.
- Deng, N. (2018) What is Temporal Ontology? *Philosophical Studies* 175(3), 793–807.
- Deutsch, D. (1997) *The Fabric of Reality: The Science of Parallel Universes and Its Implications*. Penguin.
- Duncan, M and K Miller (2014) Modal persistence and modal travel *Ratio*. 28 (3):241-255.
- Eklund, M. (2004). “Personal-identity, Concerns, and Indeterminacy.” *The Monist* 87 (4):489-511.

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

Frankfurt (2002) "Reply to John Marin Fischer." In *Contours of agency: Essays on themes from Harry Frankfurt*, ed. Sarah Buss and Lee Overton, 27-31. Cambridge, Mass.: MIT.

Gallois, A. (1994) Asymmetry in Attitudes and the Nature of Time. *Philosophical Studies* 76(1), 51-69.

Gell-Mann, M. and J. B. Hartle (1996) Classical Equations for Quantum Systems. *Physical Review D*, 47(8), 3345–3382.

Halliwell, J. J. and P. J. Dodd (2003) Decoherence and Records for the case of a Scattering Environment. *Physical Review D*, 67(10), 105018–105027.

Healey, R. (2002) Can Physics Coherently Deny the Reality of Time? In *Time, Reality & Experience*. Craig Callender (Ed.), 293–316.

Hedden, B (2015a) Options and Diachronic Tragedy. *Philosophy and Phenomenological Research*, 90(2), 423-451.

\_\_\_\_\_. (2015b) "Time-Slice Rationality." *Mind*, 124(494), 449-491.

Huggett, N. and C. Wüthrich (2013) "Emergent Spacetime and Empirical (In)coherence" *Studies in History and Philosophy of Modern Physics*, 44, 276–285.

Ismael, J. (2002) Rememberances, Mementos and Time-Capsules. In *Time, Reality & Experience*. Craig Callender (Ed.), 317–328. (Cambridge: Cambridge University Press).

Johnston, M. (1989). "Relativism and the Self". In M. Krausz (ed.), *Relativism: Interpretation and Confrontation*. Notre Dame University Press.

Joyce, R. (2001) *The Myth of Morality*. (Cambridge: Cambridge University Press).

Kovacs, D. (2016). "Self-Made people." *Mind* 125(500): 1071-1099

\_\_\_\_\_. (ms). "Diachronic Self Making".  
[https://davidmarkkovacs.weebly.com/uploads/5/5/9/5/55953415/dsv\\_-\\_website.pdf](https://davidmarkkovacs.weebly.com/uploads/5/5/9/5/55953415/dsv_-_website.pdf)

Knox, Eleanor (2013), "Effective spacetime geometry." *Studies in History and Philosophy of Physics*, 44, 346–365.

Knox, Eleanor (forthcoming), "Physical relativity from a functionalist perspective." *Studies in History and Philosophy of Modern Physics*,  
<https://doi.org/10.1016/j.shpsb.2017.09.008>.

- Lam, Vincent and Michael Esfeld (2013), “A dilemma for the emergence of spacetime in canonical quantum gravity.” *Studies in History and Philosophy of Modern Physics*, 44, 286–293.
- Lam, Vincent and Christian Wüthrich (2018), “Spacetime is as spacetime does.” arXiv:1803.04374v2.
- Le Bihan, Baptiste (2016), “Super-relationism: combining eliminativism about objects and relationism about spacetime.” *Philosophical Studies*, 173, 2151–2172.
- Le Bihan, Baptiste (2018), “Space emergence in contemporary physics: Why we do not need fundamentality, layers of reality and emergence.” *Disputatio*, 10, 71–95.
- Le Bihan, Baptiste (forthcoming), “Priority monism beyond spacetime.” *Metaphysica*.
- Maclaurin, J. and H Dyke (2002) ‘Thank Goodness That’s Over’: The Evolutionary Story. *Ratio* 15(3), 276-292.
- Maudlin, Tim (2007), “Completeness, supervenience and ontology.” *Journal of Physics A: Mathematical and Theoretical*, 40, 3151–3171.
- MacIntyre, A. (1984) *After Virtue* (Notre Dame: University of Notre Dame Press).
- McKenna (2004) "Responsibility and Globally Manipulated Agents," *Philosophical Topics*, 32: 180-84;
- McKenna (2012a) "Defending Nonhistorical Compatibilism: A Reply to Haji and Cuypers," *Philosophical Issues*, 22(1): 269-70.
- McKenna (2012b) "Moral Responsibility, Manipulation Arguments, and History: Assessing the Resilience of Nonhistorical Compatibilism," *The Journal of Ethics*, 2012(2) 160-62
- Miller, K. (2009) “Deciphering Personal-identity Conventionalism” Special Issue *Rivista di Estetica “Convenzioni”* edited by Elena Casetta and Achille Varzi 44:2 59-85
- \_\_\_\_\_. (2013) “Personal-identity minus the persons’ *Philosophical Studies*. 166(1): 91-109.
- Moss, S. (forthcoming). “Time-slice epistemology and action under indeterminacy” *Oxford Studies in Epistemology* , vol. 5.
- Ney, A. (2015) "Fundamental Physical Ontologies and the Constraint of Empirical Coherence: A Defense of Wave Function Realism" *Synthese* 192, 3105–3124.
- Nozick, Robert, (1981). *Philosophical Explanations*, Cambridge, MA: Harvard University Press.
- Oriti, Daniele (2014), “Disappearance and emergence of space and time in quantum gravity.” *Studies in History and Philosophy of Modern Physics*, 46, 186–199.
- Page, D. N. (1996) *Sensible Quantum Mechanics: Are Probabilities Only in the Mind?* *International Journal of Modern Physics D*5, 5(6), 583–596.

- \_\_\_\_\_. (2002) Mindless Sensationalism: A Quantum Framework for Consciousness. In *Consciousness: New Philosophical Perspectives*, Quentin Smith and Alexander Jokic (Eds.), 468–506. (Oxford: oxford University press).
- Prior, A. (2003) *Papers on Time and Tense (New Edition)*. (Oxford: Oxford University Press).
- Robinson, D. (2004) Failing to Agree or Failing to Disagree? Personal-identity Quasi-Relativism. *The Monist* 87(4), 512-536.
- Rovelli, C. (1991) Time in Quantum Gravity: An Hypothesis. *Physical Review D*, 43(2), 451–456.
- \_\_\_\_\_. (2004) *Quantum Gravity*. (Cambridge: Cambridge University Press).
- \_\_\_\_\_. (2007) The Disappearance of Space and Time. In *The Ontology of Spacetime*. Dennis Dieks (Ed.), 25–36. (Amsterdam: Elsevier).
- \_\_\_\_\_. (2009) Forget Time. ArXiv: 0903.3832. <http://fr.arxiv.org/abs/0903.3832>.
- \_\_\_\_\_. (2011), “A New Look at Loop Quantum Gravity.” *Classical and Quantum Gravity*, 28, 114005.
- Schechtman, M. (1996) *The Constitution of Selves*. (Ithaca: Cornell University Press).
- Shoemaker, D.W. (2003). “Caring, Identification, and Agency” *Ethics* 114 (1):88-118.
- Tallant, J. (2008) What is it to B a Relation? *Synthese* 162(1), 117-132
- \_\_\_\_\_. (2010) A Sketch of a Presentist Theory of Passage. *Erkenntnis*, 73(1), 133–140.
- \_\_\_\_\_. (forthcoming-a) Causation in a Timeless World? *Inquiry*.
- \_\_\_\_\_. (forthcoming-b) An Error in Temporal Error Theory *Journal of the American Philosophical Association*.
- Unger, P (1990). *Identity, Consciousness and Value*. (Oxford: Oxford University Press).
- Watson (1975) Free agency, *Journal of Philosophy*, 72: 205-220.
- West, C. (1996). PhD Dissertation, ANU.
- White, S. (1989). “Metapsychological Relativism and the Self” *Journal of Philosophy* 86: 298-325.
- Whiting, Jennifer, (1986). "Friends and Future Selves," *Philosophical Review* 95: 547-80.
- Wüthrich, Christian (2017), “Raiders of the lost spacetime.” In *Towards a Theory of Spacetime Theories* (D. Lehmkuhl, G. Schiemann, and E. Scholz, eds.), Birkhäuser, Basal.
- Wüthrich, Christian (forthcoming), “The emergence of space and time.” In *Routledge Handbook of Emergence* (Sophie Gibb, Robin Finlay Hendry, and Tom Lancaster, eds.), Routledge.

Yates, David (forthcoming), "Thinking about spacetime." In *Philosophy Beyond Spacetime* (Christian Wüthrich, Baptiste Le Bihan, and Nick Huggett, eds.), Oxford University Press, Oxford.

Zimmerman (1999) "Born Yesterday: Personal Autonomy for Agents without a Past," *Midwest Studies in Philosophy*.

Zimmerman, D. (2007) The Privileged Present: Defending an 'A-theory' of Time. In *Contemporary Debates in Metaphysics*, Theodore Sider, John Hawthorne and Dean Zimmerman (Eds.), 211-225. (Malden, MA: Wiley-Blackwell).