**Prudence and Perdurance**

**Abstract**

Many philosophers are sympathetic to a perdurantist view of persistence. One challenge facing this view lies in its ability to ground prudential rationality. If, as many have thought, numerical identity over time is required to ground there being *sui generis* (i.e. non-instrumental) prudential reasons, then perdurantists can appeal only to instrumental reasons. The problem is that it is hard to see how, by appealing only to instrumental reasons, the perdurantist can vindicate the axiom of prudence: the axiom that any person-stage has reason to promote the wellbeing of any other person-stage that is part of the same person as that stage. The claim that perdurantists cannot vindicate the axiom, and hence that the view should be rejected, is what we call the normative argument against perdurantism. In this paper we argue that purely instrumental rationality can ground the truth of this axiom, and hence that the normative argument against perdurantism fails.

**1. Introduction**

Many nowadays are sympathetic to a view on which persisting persons, like other persisting objects, have distinct temporal as well as spatial parts, and persist by having those temporal parts located at different times. Persons, then, are collections (or sums or aggregates) of numerically distinct entities—temporal parts, or ‘person-stages’ as they are sometimes called—united over time by relations of causal dependence and qualitative similarity. This is the view that persons persist by perduring.[[1]](#footnote-2)

Recently, there have arisen a cluster of arguments that aim to show that perdurantism is incompatible with certain obvious normative truths, and as a result should be abandoned. Many of these arguments try to show that, if true, perdurantism would lead to repugnant ethical and prudential consequences. Roughly, these arguments proceed by noting that perdurantists are committed to holding that wherever there exists a whole four-dimensional person, there also exist a plethora of shorter-lived person-like objects that are person-stages of the person (and called “subpeople” by Olson (2010), “personites” by Johnston (2017a, 2017b)). The claim is that these subpeople are intrinsically like people, and so have the same moral status as people. The problem is that, inevitably, subpeople will end up making sacrifices whose benefits they do not get to enjoy, since they no longer exist (Olson 2010; Taylor 2013; Johnston 2017a); and will be punished for actions that occurred before they came into existence (Olson 2010; Johnston 2017a, 2017b). This is morally and prudentially repugnant, and hence, the argument goes, we should reject perdurantism.[[2]](#footnote-3)

We think such arguments are misplaced. They appear to assume that persons, conceived as we will in this paper as four-dimensional perduring entities, are the fundamental seat of agency.[[3]](#footnote-4) So, it is persons who *decide* to spend three months painstakingly learning to swim in order to reap the reward of a month swimming on the barrier reef, and it is rather unfortunate subpeople who are *forced* to spend their time learning to swim *without reaping the reward of the barrier reef trip*. Thus, on this view, agency and moral patiency come apart rather oddly: subpeople are moral patients despite not being (moral) agents.

We think this is a peculiar view to take. Persons, we think, are not the seat of agency. Persons typically have massively inconsistent beliefs, desires, and preferences. Persons typically both desire that D and not D, and believe that Q, and not Q for a large range of D and Q: they do this by having temporal parts that desire D, and others that desire not D, and temporal parts that believe Q and others that believe not Q. Given this, it is hard to see how four-dimensional persons are an appropriate locus of reasons for action. What reasons for action could a four-dimensional person have, given their largely inconsistent beliefs and desires? Equally though, what we call person-slices (instantaneous temporal parts of a person) are too temporally thin to *deliberate*, or *act*, since each of these takes time.

So, it seems to us, it is short lived localised clusters of deliberating temporal parts of persons—what we will call *agent-stages*—that are the locus of reasons for action. It is these agent-stages that have reasons. It is agent-stages, then, that are the fundamental seat of agency.

In turn, if there are agent-stages who decide to learn to swim they do so *on the basis of their own reasons.* They are not hapless victims, riding the wave of choices made by the person of whom they are parts.[[4]](#footnote-5)

This picture of persons, however, raises its own problems. If agent-stages are the fundamental locus of agency, then rational requirements fundamentally fall to agent-stages not persons. The question then arises as to why such stages, who are stages of the same-person, *ought* to care about one another in a distinctively prudential manner.

It is this question that we take up in this paper. According to the normative argument against perdurantism, they have no such reason and that is why we should reject perdurantism.

We begin, in section 2, by outlining the normative argument against perdurantism. In Section 3 we hone in on the claim that we will defend in this paper. In Section 4, we describe one possible approach the perdurantist could take to vindicating the axiom of prudence: the analytic approach. While we think there is something right about some of the ideas behind this approach, we argue that it fails. In Sections 5 and 6, we present several arguments which, jointly, aim to show how the perdurantist can vindicate the axiom of prudence. Finally, in section 7, we consider a stronger axiom—the Strong Axiom of Prudence—and discuss how far the perdurantist can get in vindicating that axiom.

**2. The Normative Argument Against Perdurantism**

In what follows, we take prudential (or “selfish” or “egoistic”) concern to be that distinctive concern we have for our own wellbeing. Prudential rationality takes the familiar consequentialist instruction to promote wellbeing, but applies it within a life rather than over distinct lives. The axiom of prudence says that a person *should promote his or her own lifetime wellbeing.* That is, a person should promote their own wellbeing over the course of their life.

Theories of prudential rationality come in different varieties. They may differ regarding whether a person should *maximise or satisfice* their own wellbeing; whether a person should *consciously aim* to maximise their own wellbeing;[[5]](#footnote-6) whether a person should maximise (or satisfice) their own *objective* utility or rather their *subjective or expected* utility,[[6]](#footnote-7) in what *wellbeing* consists,[[7]](#footnote-8) and in whether they should promote the wellbeing of all of their selves equally, or are permitted to promote the wellbeing of some selves over others.[[8]](#footnote-9) Other than where explicitly stated (later in this paper) we aim to remain neutral on these issues.

We will take the axiom of prudence to be the claim that at any time t, a person, P, ought to promote P’s wellbeing at all times. So, as a first pass we will take the perdurantist constural of the axiom to be the claim that, *necessarily*, any agent-stage, P, is rationally required to promote the wellbeing of any person-stage, P\*, that is a stage of the same person as P.

Let’s unpack this idea.

First, as stated the axiom is a claim about what is necessarily the case. One could take a weaker view and think the axiom is, if true, merely contingently true: one might think that it’s true of things psychologically like us, but that there are possible creatures (creatures that nevertheless count as persons) for whom it is false. Since we are going to argue that the perdurantist can vindicate the axiom of prudence, conceived as a necessary truth, everything we say will obviously also vindicate any modally weaker version of the axiom.

Second, the rational requirement here is not an all things considered requirement. An agent-stage might have other, conflicting requirements (be they moral or prudential). Rather, we take it that the axiom says that an agent-stage has a *pro tanto* reason to promote the wellbeing of all of the person-stages that are parts of the same person as it.

Third, the axiom of prudence says that we should promote lifetime wellbeing, and so far we have interpreted this as the claim that any agent-stage has a *pro tanto* reason to promote the wellbeing of *every* person-stage that is part of the same person as that agent-stage. Let’s say that if P\* is a person-stage of the same person as P, and if P\* is later than P, then P\* is a *continuer* of P’s; and if P\* is earlier than P, then P\* is an *ancestor* of P’s. More carefully, since person-stages can overlap one another, let’s say that P\* is a continuer of P’s just in case P and P\* are parts of the same person, and P\* is at least *partly* later than P[[9]](#footnote-10); and P\* is an ancestor of P’s just in case P and P\* are parts of the same person, and P\* is at least *partly* earlier than P. As stated, then, the axiom of prudence says that every agent-stage should promote the wellbeing of each of its ancestors and continuers.

Insofar as one thinks of promoting wellbeing as *causally* promoting wellbeing, this should strike one as too strong a claim. Many worlds lack backwards causation, and in such cases agent-stages will be unable to causally promote the wellbeing of any (let alone all) their ancestors.

Now, on some views of wellbeing it is possible for agent-stages to promote the wellbeing of some of their ancestors in virtue of satisfying those ancestors’ desires.[[10]](#footnote-11) But even if one accepts this kind of view (which is in itself controversial) it will not always be possible for an agent-stage to promote the wellbeing of an ancestor: namely in cases in which that ancestor has no desires that the agent-stage in question can satisfy.

Given that agent-stages can only be rationally required to do what they can do, and given that we want to (largely) remain neutral about theories of wellbeing, we will suppose that the sense in which persons should promote their own lifetime wellbeing is the sense in which, at any time, a person should promote the wellbeing of itself at all times to which it has causal access. Thus, in worlds without backwards causation this amounts to saying that any agent-stage has a *pro tanto* reason to promote the wellbeing of all of its *continuers*. This is the version of the axiom that we will focus on, though everything we say can be applied to any ancestors an agent-stage has, whose wellbeing can be causally promoted.

Fourth, what is it to promote wellbeing? Recall that we are remaining neutral on what wellbeing consists in. Still, we want to distinguish what we might call *partial* wellbeing from *overall* wellbeing. To get a handle on this distinction let’s imagine for a moment that wellbeing consists in being healthy, wealthy, wise and attractive. Then an agent-stage partially promotes wellbeing if it promotes *any* of these aspects of wellbeing. It promotes overall wellbeing if it promotes *all* of these aspects of wellbeing.[[11]](#footnote-12) In what follows when we talk about promoting wellbeing we mean partially promoting wellbeing unless we say otherwise.

So, as we will interpret the axiom of prudence, it says the following:

**Perdurantist Axiom of Prudence**: necessarily, any agent-stage is rationally required to partially promote the wellbeing of each of its continuers.

A stronger axiom of prudence says the following:

**Strong Perdurantist Axiom of Prudence**: necessarily any agent-stage is rationally required to promote the *overall* wellbeing of each of its continuers.

We will consider the stronger version of the axiom in Section 7, but for the remainder of the paper we focus only on the weaker axiom. That is because even the weaker axiom is imperilled by the normative argument against perdurantism.

To see why, notice that there are broadly two accounts of what grounds the truth of the axiom of prudence. According to Humean or instrumental views of rationality, what grounds its being true that at any time t, a person P ought to promote P’s wellbeing at all times, are P’s desires.[[12]](#footnote-13) *Strict* views of instrumental rationality will say that it is P’s actual desires, while *idealised* views of instrumental rationality will say that it is P’s counterfactual desires, in particular, it is usually thought to be P’s (subjectively) ideal desires—i.e., what P would want for herself for its own sake (i.e. non-instrumentally) if P was in (what P takes to be) more ideal conditions of reflection.[[13]](#footnote-14) For example, suppose that, being unaware of the joys that canine companionship would bring me, I currently have no desire to acquire a pet dog. However, if I fully and accurately foresaw (both intellectually and emotionally) the delights of canine companionship, I would want to acquire a canine companion for myself. According to an idealised view of instrumental rationality, although acquring a canine companion does not satisfy any actual desire of mine, I nevertheless have prudential reason to acquire a canine companion (and hence to form a desire to acquire one) inasmuch as under these more ideal conditions of reflection—say, where I possessed fuller knowledge of the joys of canine companionship—I would desire this for myself.[[14]](#footnote-15)

By contrast, according to non-Humean/non-instrumental views of rationality there exist *sui generis* *prudential reasons*: reasons that are not moral reasons, nor reasons arising from P’s (actual or ideal) desires, and these ground the truth of the axiom of prudence.[[15]](#footnote-16)

Many philosophers have thought that instrumental rationality, in either its strict or idealised versions, cannot ground the truth of the axiom of prudence, because there are possible agent-stages who have a pattern of desires which fails to make it true that they should promote the wellbeing of all of their continuers.

Consider an agent-stage of Jeremy at aged 13. One can imagine that this stage, Jeremy-at-13, has a pattern of desires that do not ground his having reason to promote the wellbeing of a person-stage of Jeremy at 89. We can imagine that Jeremy-at-13 simply does not care what happens to him in old age. So, Jeremy-at-13’s actual desires do not ground his having reason to promote the wellbeing of Jeremy-at-89. Likewise, we seem able to imagine that Jeremy-at-13 is such that even his subjectively idealised desires are such that they fail to give him reason to promote the wellbeing of Jeremy-at-89. In that case, if there are only instrumental reasons it follows that Jeremy-at-13 *has* no reason to promote the wellbeing of Jeremy-at-89 and hence that the axiom of prudence is false.[[16]](#footnote-17)

This has prompted many to conclude that the axiom of prudence must be grounded in *sui generis* prudential reasons. This has been taken to present special problems for the perdurantist. For, the thought goes, the only thing that could ground there being such *sui generis* prudential reasons is numerical identity. The reason that P at t should care about P at all future times in this distinctively prudential manner—and should do so regardless of P’s own desires—is that P at t is numerically identical with P at each of these times. P should care about *itself*, and so the obtaining of numerical identity over time grounds the obtaining of *sui generis* prudential reasons.

By contrast, non-identity relations do not seem able to ground there being such reasons. For the sorts of non-identity relations that obtain between P at one time and P at another time are similar to those that obtain between P at one time and Q at some other time. Yet we do not want to say that P has *prudential* reason to care about Q. The problem facing perdurantism, then, is that because the view renders the connection between stages of the *same* person as similar in kind to the connection between stages of *different* persons, it undermines there being any *sui generis* prudential reasons.

Here is Sidgwick back in 1884:

Grant that the Ego is merely a system of coherent phenomena, that the permanent identical ‘I’ is not a fact but a fiction, as Hume and his followers maintain; why, then, should one part of the series of feelings into which the Ego is resolved be concerned with another part of the same series, any more than with any other series? (418-19).[[17]](#footnote-18)

Schechtman expresses similar sentiments when she notes that self-interested concern is an emotion that is appropriately felt only toward *my own self and not toward someone like m*e (Schechtman 1996, 52).[[18]](#footnote-19) She notes that it “doesn’t matter if the person who gets my paycheck is more *like* me than someone else; I am only compensated if *I* get the money” (Schechtman 1996, 52-3).[[19]](#footnote-20), [[20]](#footnote-21)

So, the problem for perdurantists who, like us, take agent-stages (rather than whole four-dimensional persons) to be the fundamental locus of agency, is that it is not possible to appeal to numerical identity to ground there being *sui generis* prudential reasons, since the agent-stages of persons are numerically distinct from one another. But, as we noted earlier, it is these agent-stages that are the fundamental locus of agency, and hence of rational requirement, even if persons themselves are four-dimensional objects.

Thus, the perdurantist is faced with the normative argument against perdurantism.[[21]](#footnote-22) That argument proceeds as follows. The axiom of prudence is true. But the axiom can only be true if P at one time has *sui generis* prudential reasons to promote the wellbeing of P at future times. That is because mere instrumental reasons cannot ground the truth of the axiom of prudence. P at one time, however, can only have *sui generis* prudential reasons to promote the wellbeing of P at any future time if P at t is numerically identical with P at each of those future times. But, according to perdurantism, P at t is numerically distinct from P at each of these times. Hence, perdurantism is incompatible with the truth of the axiom of prudence, and since the axiom is true, perdurantism is false.

In fact, we do not endorse the idea that *sui generis* prudential reasons must be grounded by numerical identity. Hence, we do not accept that the perdurantist cannot appeal to such reasons.[[22]](#footnote-23)In this paper, however, we are primarily interested in showing that the perdurantist axiom of prudence can be vindicated by appealing only to instrumental reasons.

**3. Prudence and Humeanism**

We will say that P has a *strict instrumental reason* to promote the wellbeing of P\* iff (a) P and P\* are temporal parts of the same person and (b) promoting the wellbeing of P\* is a way to satisfy P’s actual desires. Another way to put this is to say that P has an instrumental reason to promote the wellbeing of P\*, iff P and P\* are temporal parts of the same person and P’s total set of non-instrumental desires, and what is logically and instrumentally entailed by those desires in combination with the non-normative facts, grounds P having a reason to promote P\*’s wellbeing.

We will say that P has an *idealised instrumental reason* to promote the wellbeing of P\* iff (a) P and P\* are parts of the same person and (b) promoting the wellbeing of P\* is a way to satisfy P’s subjectively idealised desires. Another way to put this is to say that P has an instrumental reason to promote the wellbeing of P\*, iff P and P\* are temporal parts of the same person, and P’s total set of subjectively idealised non-instrumental desires, and what is logically and instrumentally entailed by those desires in combination with the non-normative facts, grounds P having a reason to promote P\*’s wellbeing.

Then we will say that P has an instrumental reason to promote the wellbeing of P\* iff P either has a strict or an idealised instrumental reason to do so.

So, for instance, in the most obvious case P has an instrumental reason to promote the wellbeing of its continuers when P has a desire (instrumental or otherwise) to do so. But it need not be that P’s reasons to promote the wellbeing of some particular person-stage, P\*, issues from P’s desiring to promote P\*’s wellbeing. It could be that P has other desires whose satisfaction gives P instrumental reason to promote P\*’s wellbeing.

The problem with grounding the perdurantist axiom of prudence in instrumental reasons is that it seems obvious that there are possible agent-stages, P and P\*, such P\* is P’s continuer, and yet P’s desires provide P with no reason to promote P\*’s wellbeing. This is what we call the Possibility Claim:

**Possibility Claim:** It is (metaphysically) possible that there are agent-stages P and P\* such that P\* is P’s continuer, and promoting P\*’s wellbeing is not a way to satisfy P’s actual/subjectively idealised desires.

The Possibility Claim seems very plausible. If it’s true, though, the truth of the axiom of prudence cannot be grounded in instrumental reasons, and hence if those are the only reasons to which the perdurantist can appeal, it follows that if the axiom is true, then perdurantism is at the very least actually false (and with not much argument one might try to show that it is necessarily false).

In fact, you might even think that a much weaker claim is also true. Call this the Psychological Claim.

**Psychological Claim:** It is psychologically possible that there are agent-stages P and P\*, that that P\* is P’s continuer, and promoting P\*’s wellbeing is not a way to satisfy P’s actual/subjectively idealised desires.

If the Psychological Claim is true then the perdurantist cannot even retreat to a weaker reading of the axiom of prudence on which it is, if true, psychologically necessary (say).

In what follows we focus on the Possibility Claim, since we are interested in the more modally robust version of the axiom.

We want to distinguish two possibilities, each of which would render the Possibility Claim true. We distinguish these because different resources need to be brought to bear to show why neither is in fact possible. To do so, we introduce the idea of an *adjacent agent-stage.*

Consider an agent-stage, P, which exists throughout a short temporal interval T, where T is composed of instants T1 to T10. P\* is an agent-stage that exists throughout the interval composed of T11 to T21. Call P\*, P’s *adjacent agent-stage continuer.* P\* is, at it were, the ‘very next’ agent-stage continuer of P’s which does not overlap any of P itself.[[23]](#footnote-24)

Then the Possibility claim will be true if it is possible that there is an agent-stage that has an adjacent agent-stage continuer such that that stage’s desires do not provide it with reason to promote the wellbeing of that adjacent continuer. We will call this the Adjacent-Stage Possibility Claim.

**Adjacent-Stage Possibility Claim:** it is possible that there is an agent-stage P such that P\* is P’s adjacent continuer, and P’s actual/idealised desires do not provide P with any reason to promote the wellbeing of P\*.

The Possibility claim will also be true if it is possible that there is an agent-stage that has a continuer that is more temporally distant than an adjacent continuer, and where that stage’s desires do not provide it with a reason to promote the wellbeing of that more temporally distant stage.

**Non-Adjacent-Stage Possibility Claim:** it is possible that there is an agent-stage P such that (i) P\* is P’s continuer, (ii) P\* does not overlap P’s adjacent continuer, and (iii) P’s actual/idealised desires do not provide P with any reason to promote the wellbeing of P\*.

Non-adjacent agent-stages, then, are continuer stages that are entirely distinct from adjacent agent-stages (i.e., they do not overlap adjacent agent-stages). Some such stages might be relatively temporally near to the agent-stage in question, while others might be temporally quite far (as when we consider a particular agent-stage of Jeremy at 13, and another of Jeremy at 89).

Our aim will be first to show that the Adjacent-Stage Possibility claim is false, and, from there, to show that the Non-Adjacent-Stage Possibility Claim is false. First, however, we turn to consider a view that promises to straightforwardly show that the Possibility Claim is false. Ultimately, we will argue that this view fails—or, at least, that the version that succeeds may not be very attractive. But we think that some of the lessons learned here will be useful later when we come to present our own arguments against the Possibility Claim.

**4. The Analytic Approach**

According to one view of persons, person-stages count as stages of a persisting person only insofar as they desire to promote the wellbeing of the person of which they are stages. Very roughly, on such views it is partly *constitutive* of P\* being a continuer or ancestor of P’s, that P directs certain attitudes towards P\*: in particular, that P directs attitudes of prudential concern towards P\* (that is, P desires to promote P\*’s wellbeing). Thus, on this view it is analytic that P and P\* are agent-stages of the same person if and only if P and P\* desire to promote each other’s wellbeing in a distinctively prudential sort of way. So, on this view, it is constitutive of P\* being a continuer or ancestor of P that P have the right sorts of desires to ground P’s having reason to promote P\*’s wellbeing, and hence the Possibility Claim is false.

Whiting (1986) defends a view like this. More generally, *conventionalists*[[24]](#footnote-25) about personal-identity hold that the personal-identity relation—that is, the relation that obtains between person-stages that are stages of the same person—is in part determined by the obtaining of certain kinds of attitudes of prudential concern between the stages in question.

We are broadly sympathetic to *something* like this view. But there is a worry that a version of conventionalism that can resist the Possibility Claim in this manner is not very plausible. For it will turn out that if person-stage P does *not* care prudentially about P\*, then regardless of which other relations obtain between P and P\*, P\* is not an ancestor or a continuer of P\*. So, if an agent-stage of 13 year old Jeremy does not care, prudentially, about an agent-stage of 89 year old Jeremy, then it just turns out that the latter is not a continuer of the former. There are ‘two Jeremys’, one who begins life as a child and ceases to exist before old-age, and another who comes into existence later in life and lives to be old. But, you might think, this is not overly plausible. That is why many versions of conventionalism appeal to attitudes of prudential care in a more subtle manner than this.

For instance, some versions of conventionalism hold that what determines which continuers are P’s are the overlapping chains of (*inter alia)* attitudes of self-concern emanating from P. On this kind of view, it is constitutive of P\* being an adjacent agent-stage continuer of P’s that P direct certain sorts of prudential attitudes towards P\*. So, this kind of view is one on which the Adjacent-Stage Possibility Claim is false. Necessarily, if P\* is an adjacent agent-stage continuer of P’s, then P is rationally required to promote the wellbeing of P\*. What this view does not seem able to do (at least by appealing directly to these attitudes) is to make false the Non-Adjacent-Stage Possibility Claim. For we can stipulate that there are such chains connecting an agent-stage of Jeremy at 13 with an agent-stage of Jeremy at 89. Then according to this view these two agent-stages are stages of the same person. Yet the former stage does not care prudentially about the latter. So if it is just the presence of these attitudes that grounds the truth of the axiom of prudence, then the axiom is false.

Perhaps a version of this view could appeal to the idealised desires of agent-stages, rather than just their actual attitudes, in order to resist the Non-Adjacent-Stage Possibility Claim. For now, we simply note that, on its face, this version of conventionalism will not do.

Still other versions of conventionalism hold that what matters is not to which *particular* stages P directs its attitudes of prudential concern; but, rather, around which relation P ‘organises’ those attitudes.[[25]](#footnote-26) If, for instance, P *typically* directs prudential concern only towards stages with which P is psychologically continuous, then all and only stages that are psychologically continuous with P are its ancestors or continuers.

This view, however, is one that leaves us with no obvious way to resist either the Adjacent-Stage or Non-Adjacent-Stage Possibility Claims. To see this, suppose that P organises its practises around the relation of psychological continuity. If this is so then, for the most part, P cares prudentially about stages that are connected to P via psychological continuity. Still, this does not entail that P cares prudentially about *all* such stages. It is consistent with the view that there are both adjacent and non-adjacent stages that P does not prudentially care about. Then P’s actual attitudes of prudential concern do not ground its being rationally required for P to promote the wellbeing of these stages, be they adjacent or non-adjacent.

The general problem here is that the more plausible the version of conventionalism as an account of personal identity, the more likely it is that there are possible cases in which P\* counts as being a continuer of P’s, and yet P has (actual) prudential attitudes that do not ground P’s having a reason to promote P\*’s wellbeing.

In what follows we begin by trying to show that the Adjacent-Stage Possibility Claim is false. We will then use this to argue that the Non-Adjacent-Stage Possibility Claim is false.

**5. Against The Adjacent-Stage Possibility Claim**

In what follows we present two arguments against the Adjacent-Stage Possibility Claim. The first of these is inspired by the conventionalist idea that it is constitutive of P\*’s being an adjacent continuer of P’s that P has certain desires, which ground P’s having reason to promote P\*’s wellbeing.

Here is the idea. We take it to be uncontroversial that, necessarily, an agent-stage has instrumental reason to promote the wellbeing of agent-stages that satisfy that stage’s actual desires. Call the stage that satisfies an agent-stage’s desire the *desire satisfying stage.* So, where P\* satisfies P’s desire, P has reason to promote P\*’s wellbeing. (Remember that we are only talking about the promotion of wellbeing here, not the promotion of *overall* wellbeing. It may be that the fact that P\* satisfies that desire does not give P a reason to promote P\*’s overall wellbeing, for perhaps there are aspects of P\*’s wellbeing that are irrelevant to its capacity to satisfy that desire. We return to think about overall wellbeing in Section 7).

We will argue that necessarily, for any agent-stage, if that agent-stage has an adjacent agent-stage then that adjacent-stage continuer satisfies at least some of its desires. If this is true, then it entails that the Adjacent-Stage Possibility Claim is false.

Why think that, necessarily, for any agent-stage, if that agent-stage has an adjacent agent-stage continuer then that adjacent-stage continuer satisfies at least some of its desires? The truth of this claim hinges on a further claim about what it is to be an agent-stage. Here is where the idea is loosely inspired by conventionalist views of personal-identity. We think it is *constitutive* of being an agent-stage that agent-stages have *future-directed desires*. We will say that an agent-stage P has a future-directed *desire* just in case P has a desire that can only be satisfied by some continuer of P’s.[[26]](#footnote-27)

Something that lacked these would, we say, fail to be an agent-stage. Agent-stages are, of necessity, things that deliberate, intend, and act. As such, they will, of necessity, have future-directed desires. Moreover, we think it is necessary that at least some of an agent-stages’ future directed desires are satisfied by their adjacent agent-stages.

Now, you might think that there are possible agent-stages whose only future-directed desires are ones that are satisfied by more distant (non-adjacent) agent-stages. We don’t think that such agent-stages are possible. To be sure, an agent-stage can have a single non-instrumental desire, such as the desire to play a symphony in 10-years’ time. And that non-instrumental desire can only be satisfied by a non-adjacent stage. But that desire must give rise to a host of instrumental desires, and those instrumental desires will be such that at least some of them are satisfied by adjacent agent-stages. For in order to have desires about how things go at later times, an agent-stage must have at least *some* desires for how things go at temporally near times. For instance, suppose that a current agent-stage only non-instrumentally cares about performing a piano symphony at the royal gala in three years time. She has no non-instrumental desires about what happens in the interim. Still, we say, she must have *some* instrumental desires about what happens in the interim: she must at the very least desire to *survive* during the intervening period, which entails having desires to eat, move, and so on. And at least some of those desires are ones that can only be satisfied by an adjacent stage (for instance, the adjacent agent-stage must survive if later ones are to survive).

Now, perhaps as stated this is not quite right. Suppose that an agent-stage falsely believes that it will not persist beyond two hours time. Now consider the agent-stage that exists just prior to the moment it *believes* death will come. Perhaps that agent-stage has no forwad-looking non-instrumental desires, and hence no desires that could ground it having reason to promote the wellbeing of its adjacent agent-stage continuer.

The tempting thing to say here is that the agent-stage in question nevertheles has *conditional* desires: were it to believe that it will survive (as in fact it will), it *would have* various desires that would ground it having such reasons. (In particular, one might think that it’s the desires that an agent-stage would have, if it had true beliefs, that are relevant here). One might worry, though, that appealing to conditional desires takes us away from a view on which it is our actual desires that ground our instrumental desires. After all, conditional desires are just the desires we *would* have, if we had different beliefs. So it’s worth noting that in appealing to conditional desires, the resulting view comes a little closer to the view that it is our idealised non-instrumental desires that ground our instrumental reasons. Having said that, it’s still the way the agent-stage is that grounds her having the reasons she does: after all, she really is such that she is disposed to have that desire, conditional on believing that she will live past two hours.

To put the point a somewhat different way, most accounts of personal-identity are ones on which certain kinds of causal connections between person-stages are necessary for those stages to be parts of the same person.[[27]](#footnote-28) So, part of what makes it the case that agent-stage P\* is the (future) adjacent continuer of agent-stage P, is that P’s desires are causally efficacious in bringing about P\*’s actions. For instance, part of what makes it the case that P\* is P’s adjacent continuer is that P’s tiredness, and desires to sit, causes it to be the case that P\* sits down and takes a rest.[[28]](#footnote-29) On this picture, the fact that P\* satisfies P’s desires is (in part) what makes it the case that P\* is P’s adjacent continuer. So, we say, it is of the nature of agent-stages that they have future-directed desires, and, in part, what makes it the case that P\* is an adjacent continuer of P, is that (some of) P’s immediate desires are satisfied by P\*.

In turn, it follows that, necessarily, for any agent-stage P, P will have some desires that are satisfied (and can only be satisfied by) P’s adjacent continuer. These desires, in turn, ground its being the case that, necessarily, P has reason to promote the wellbeing of its adjacent continuer. Hence the Adjacent-Stage Possibility Claim is false.

That is our first argument against the Adjacent-Stage Possibility Claim. Here is our second. So far, we have only appealed to the actual desires of agent-stages. In what follows we move to appealing to their subjectively idealised desires. Here is the idea. We’ll begin by introducing the *principle of reflection for desires.* That principle is analogous to an amended version of the principle of reflection for beliefs, first introduced by van Fraassen (1984). Reflection for beliefs says that you should defer to the beliefs that you anticipate having in the future: that is, if you’re going to come to believe Q (or have credence z in Q) then you should believe that Q (or have credence z in Q) now. (Importantly, Q is a proposition, and hence is taken to be fully specific. So, if at t Jeremy believes that it is raining, then what Jeremy believes is that *it is raining at t*. This matters; if the content of Jeremy’s belief, at t, were simply *it is raining* then it needn’t automatically be rational for Jeremy, at some earlier time, to believe *that* proposition. After all, at the earlier time it may not be raining. So, belief reflection is the claim that if Jeremy now believes that he will believe *that it is raining at t,* then he should, now, believe *that it is raining at t.*) Modified versions of reflection for belief say that you should defer to the beliefs that you anticipate having in the future, unless you believe that in the future you will be irrational or will have lost evidence. Modified versions of belief reflection are widely endorsed, because agents for whom (modified) belief reflection is false can be turned into money pumps.[[29]](#footnote-30)

Similar considerations motivate modified desire reflection.[[30]](#footnote-31) Modified desire reflection says that you should defer to the desires that you anticipate having in the future. That is:

**Modified Reflection for Desires**: If you believe, or anticipate, that you will come to desire Q, then you should desire that Q now, as long as you do not anticipate that in the future you will be irrational or will have lost evidence.[[31]](#footnote-32)

(Again, it’s important here that Q is a maximally specific proposition. Consider Jeremy again. Suppose he’s just eaten a huge breakfast and is completely full, but he believes that by noon he will desire smashed avocado on toast. Clearly Jeremy does not, now, have reason to desire the smashed avocado on toast (given that he is full). Rather, modified desire reflection says that, given that Jeremy believes he will desire smashed avocado on toast at noon, he should now desire smashed avocado on toast at noon).

Agents for whom modified desire reflection is false can be turned into money pumps. Suppose that Freddie’s default dinner meal is beef. Now suppose that as of today (Monday) Freddie would prefer kangaroo to beef for dinner in two-day’s time (Wednesday). Suppose he thinks it is worth $1.00 to make sure that he gets the kangaroo not the beef on Wednesday. Now suppose Freddie also believes that tomorrow, he will prefer that on Wednesday he will have beef rather than kangaroo. Tomorrow, he will pay $1.00 to get the beef not the roo on Wednesday. Then Freddie can end up paying $2.00 to gain nothing.

Modified Reflection for Desires talks about what *you* believe or anticipate that *you* will come to desire. Since we are taking agent-stages to be the fundamental locus of agency, it’s perhaps not true that an agent-stage believes that, or anticipates that, *it* will have certain desires, since it may not be around to have those desires. Nevertheless, since surely agent-stages do anticipate things being thus and so even after they have ceased to exist, we can amend reflection for desires to say the following:

**New Modified Reflection for Desires**: If agent-stage P anticipates[[32]](#footnote-33) preferring that x be true, then P should now prefer that x be true as long as P does not think that this preference will be the product of irrationality or lost evidence.

Of course, modified reflection for desires is controversial.[[33]](#footnote-34) And, you might think, it’s not a principle that any instrumentalist about rationality is likely to endorse, at least as a fully general claim about what agent-stages should prefer. For if we are instrumentalists about rationality then the only grounds for the truth of such a principle will be agent-stages’ actual or idealized desires. And perhaps some agent-stages have actual or idealized desires that fail to give them a reason to accept reflection for desires.

Certainly, we think that an appeal only to the actual desires of agent-stages cannot ground all agent-stages having reason to accept reflection for desires. Nor can we hope to completely convince you that the idealized desires of any possible agent-stage will ground the truth of that principle for that agent-stage. In fact, though, we think that insofar as the principle is plausible at all, it’s plausible for all possible agent-stages.

One might initially think that only agent-stages who care about becoming money pumps will have reason to accept reflection for desires, and that some agent-stages won’t care about this. But, arguably at least, agents for whom desire reflection is false can be turned into pumps for *whatever it is that they do care about*. Suppose that Freddie has three non-instrumental desires: a desire for roo, beef, and running. Suppose that as of today (Monday) Freddie would prefer kangaroo to beef for dinner in two-day’s time (Wednesday). Then it is not so implausible to think that there is *something* that is valuable to Freddie (even if only a very small amount of that thing) that he will trade to make this so. Suppose that Freddie is willing to trade 50 seconds of running to bring it about that he gets the kangaroo not the beef on Wednesday. Now suppose Freddie also believes that tomorrow, he will prefer that on Wednesday he will have beef rather than kangaroo. Tomorrow, he will trade 50 seconds of running to get the beef not the roo on Wednesday. Then Freddie can end up having forgone 100 seconds of running to be no better off at all. Overall, he will be worse off with regard to satisfying his desires, since he desires to run.

As long as there is something of value that Freddie is prepared to trade in order to bring about his preference, he can be turned into a pump for that thing (in this case running). Since Freddie values the thing he trades, he surely will have idealised desires not to be pumped of that thing, and hence his idealised desires will ground his having reason to accept reflection for desires. What is true of Freddie is true for us all.[[34]](#footnote-35)

So, suppose that reflection for desires is true. Then consider P, and P’s adjacent continuer, P\*. Let’s stipulate that P\* desires its own wellbeing. We will return to this stipulation shortly. Then New Modified Reflection for Desires says that insofar as P anticipates having these desires—or anticipates preferring that things are thus and so—then P ought, now, to have those desires/preferences. For now, let’s make another supposition: that P *does* anticipate having the desires/preferences in question. Then it follows by New Modified Reflection for Desires that P *ought* to have certain desires, now, regarding how things are for P\*. The idealized instrumentalist about rationality can then say that it is the desires that P ought to have which ground P’s having reason to promote P\*’s wellbeing.

So, if New Modified Reflection for Desires is true for all agent-stages, then it follows that, necessarily, for any agent-stage P and adjacent stage P\*, if P\* desires its own wellbeing, and if P anticipates having (certain of) P\*’s desires/preferences, then P has a reason to promote P\*’s wellbeing.

This is all well and good. But it only follows that the Adjacent Possibility Claim is false if it’s both the case that: (a) necessarily, for any agent-stage P and adjacent agent-stage P\*, P anticipates having (certain of) P\*’s desires/preferences; and (b) necessarily, any agent-stage P\* desires its own wellbeing. But these are both surely very controversial claims. Let’s start with (a). We will argue that (a) is true.

Everything we have said so far can be used to argue for the claim that, necessarily, for any agent-stage P and adjacent agent-stage P\*, P anticipates having (certain of) P\*’s desires/preferences. We’ve already argued for a view on which P\* is an adjacent agent-stage continuer of P’s in part because P\* fulfils/carries through P’s immediate desires and intentions. Part of what it is for P to have those immediate desires and intentions is to anticipate their satisfaction/fulfillment. It’s hard to see what sense can be made of the idea of P intending to, say, sit down, without its also being the case that P *anticipates* sitting down. And perhaps it’s also hard to see what can made of the idea of P having an immediate desire to sit down, without its being the case that P anticipates sitting down. So, we think it plausible that it is constitutive of P\* being P’s adjacent agent-stage continuer that P anticipates at least some of P\*’s experiences, desires, and preferences.

But what of (b), the claim that, necessarily, any agent-stage P\* desires its own wellbeing? That claim, one might think, is surely false. And if an agent-stage only has reason to promote a continuer stage’s wellbeing conditional on that continuer itself desiring to do so, then sometimes agent-stages won’t have such reason (at least, not arising from reflection).

Now, the perdurantist might, at this stage, say that she was concerned to show that there is no particular *diachronic* problem of prudence arising from endorsing perdurantism. She might point out that if we endorse reflection, then it follows that insofar as the person-stages of a person are each concerned with their *own* wellbeing, they thereby have reason to be concerned with *one another’s* wellbeing. And this would still, we think, reflect signifant inroads here.

Nevetheless, the axiom of prudence does not merely say that a person-stage has reason to promote the wellbeing of its continuer stages conditional on those stages desiring their own wellbeing. So the question arises as to whether the perdurantist can do better here. We think she can, but only if she no longer remains neutral regarding a theory of wellbeing.

Suppose the perdurantist were to adopt a desire satisfaction theory of wellbeing, according to which wellbeing consists in the satisfaction of desires.[[35]](#footnote-36) Now, to be clear, we are not going to defend this view of wellbeing here. Rather, our aim is just to show that if one endorses such a view, then one can show that necessarily, any agent-stage P\* desires its own wellbeing. Moreover, insofar as the desire fulfillment theory has a claim to be “the dominant view of welfare among economists, social-scientists and philosophers, both utilitarian and non-utilitarian”[[36]](#footnote-37), this argument may have wide appeal.

How does the satisfaction of desires contribute to wellbeing, and to whose wellbeing does it contribute? Well, suppose that the *time-of-desiring* is the time at which a desire is held. Suppose the *time-of-satisfaction* is the time at which a desire is satisfied. Suppose the *time-of-wellbeing* is the time at which a person-stage is made better or worse off by the desire being satisfied.

According to *concurrentism* a desire is not satisfied when the time-of-desiring and the time-of-satisfaction come apart and so no wellbeing accrues. According to the *time-of-satisfaction view*, the person-stage who benefits from a desire being satisfied is the person-stage that exists at the time-of-satisfaction. So on this view, if P1 desires to sit down and it is P2 who sits, then it is P2 whose wellbeing is promoted. Finally, on the *time-of-desiring view,* the person-stage who benefits from the desire being satisfied is the person-stage whose desire it is that is satisfied.[[37]](#footnote-38) So on this view, since it is P1 who desires to sit, and P2 who satisifes this desire, it is P1 whose wellbeing is promoted.

Suppose one endorses the time-of-desiring view. Now suppose that P\* is P’s adjacent continuer, that P\* has a bunch of desires, and that satisfying those desires constitutes promoting P\*’s wellbeing. Consider one such desire, D. According to reflection, P should desire D. Now, it is likely that P cannot itself satisfy D (assuming that D is a future-directed desire of P\*’s, and not a past-directed one). Still, even if D is a future-directed desire of P\*’s, P can promote the satisfaction of D. P can act, and plan, in such a way that the satisfaction of P is more, or less, likely. And since P desires D, P clearly has reason to promote the satisfaction of D. But notice that promoting the satisfaction of D just is promoting P\*’s wellbeing, since satisfying D just is (partially) promoting P\*’s wellbeing according to the desire satisfaction theory.

Putting all this together, then, necessarily, P\* is P’s adjacent agent-stage continuer only if P anticipates at least some of P\*’s experiences, desires, and preferences. P’s desires, whatever they are, ground P’s having reason to endorse modified desire reflection. So, if P anticipates at least some of P\*’s experiences, desires, and preferences, then P ought, now, to have those desires (by desire reflection). If P ought, now, to have those desires, then this is to say that P subjectively ideally has those desires. P’s subjectively ideally having those desires amounts to P’s subjectively ideally desiring that P\*’s desires are satisfied (since P’s subjectively idealized desires just are P\*’s desires, and P of course desires that it’s own desires are satisfied). This gives P reason to promote the satisfaction of those desires. But by promoting the satisfaction of those desires, P just is promoting P\*’s wellbeing, since P\*’s wellbeing just consists in its desires being satisfied. So, necessarily, if P\* is P’s adjacent continuer, then P has a reason to promote P\*’s wellbeing.

Thus, we have provided two arguments for the conclusion that the Adjacent-Stage Possibility Claim is false.

To be clear, we don’t think these arguments are unassailable. One might deny that reflection for desires is true. Or one might think it’s true for most agents, but that there are at least *possible* agents whose desires do not ground that principle’s truth. One might deny that, necessarily, agent-stages anticipate (at least some) of the desires/experiences of their adjacent agent-stages. Or you might deny that desire satisfaction theories are the right accounts of wellbeing. We think these claims are plausible, but others may not.

If you deny one or more of the above claims then it will not be possible to vindicate the perdurantist axiom of prudence by appealing only to instrumental reason. Still, the perdurantist who takes this route might argue that, although there are possible agent-stages

for whom reflection is false, or possible agent-stages that do not anticipate any of the experiences/desires of their adjacent continuer; nevertheless, such agent-stages will be psychologically *very* different indeed from you and me. The perdurantist might then take the view that although the axiom of prudence is not a necessary truth, it is nevertheless true of all persons/agents that are anything at all like you and me: it is at least psychologically necessary. So, she might say, she can vindicate a modally weaker version of the axiom of prudence and, she might try to argue, that is sufficient. Indeed, she might suggest that it is unclear whether agent-stages that are so radically different from you and me really do have a *pro tanto* reason to promote the wellbeing of their continuers.

This is not the view we take. In what follows, we will proceed on the assumption that the

Adjacent-Stage Possibility Claim is false, and use that to argue for the falsity of the Non-Adjacent Possibility Claim.

**6. The Non-Adjacent Possibility Claim**

Let’s suppose that P1 through P4 are agent-stages of the same person, and that P2 is the adjacent continuer of P1, and P3 is the adjacent continuer of P2, and so on. If the Adjacent-Stage Possibility claim is false, it follows that P1 has a reason to promote the wellbeing of P2, and P2 has a reason to promote the wellbeing of P3, and P3 has a reason to promote the wellbeing of P4. But, of course, it does not follow from this that P1 has a reason to promote the wellbeing of P3 or P4.

In what follows we offer two arguments that aim to show that if the Adjacent-Stage Possibility Claim is false, then so too is the Non-Adjacent-Stage Possibility Claim: if P1 has reason to promote P2’s wellbeing, and P2 has reason to promote P3’s wellbeing, then P1 has reason to promote P3’s wellbeing (and so on).

Here is the first such argument.

Recall that part of the story we told about why P1 has reason to promote P2’s wellbeing is that P1 anticipates at least some of P2’s experiences and desires because, necessarily, any agent-stage anticipates at least some of its adjacent continuer’s experiences and desires. Given this, we know that P2 anticipates at least some of P3’s experiences and desires. Let’s take the limiting case of anticipation, in which P1 simply anticipates “being in P2’s position” in the sense of anticipating all of P2’s experiences and desires. Suppose P2 is similarly related to P3 (and so on down the line). Then we could say that P1 anticipates being P2, and that P2 anticipates being P3. Now suppose we put P1 in an epistemic position in which P1 knows these facts. What desires can we expect P1 to form? Well, P1 will surely desire *not* to *anticipate anticipating* unpleasant events. That is P1 has subjectively idealised desires not only that things go well for P2, but also that things are such that P2 anticipates things going well for P3. That is, P has subjectively idealised desires not to anticipate anticipating unpleasant events. That being so, P1 has subjectively idealised desires that ground its having reason to promote P3’s wellbeing. And, *mutatis mutandis,* for P4.

Suppose, though, that P1 does not anticipate *all* P2’s experiences/desires, and likewise for P2 with respect to P3. Then it might be that although P1 anticipates certain of P2’s experiences/desires, and P2 anticipates certain of P3’s experiences/desires, P1 does not anticipate anticipating *any* of P3’s experiences/desires (for the experiences/desires of P2’s that P1 anticipates are not ones that include an anticipation of P3’s experiences/desires). But then P1’s desires do not ground P1’s having reason to promote P3’s wellbeing. And if so, there’s a possible agent-stage (P1) whose desires do not ground it having reason to promote the wellbeing of all of its non-adjacent continuers: the Non-Adjacent-Stage Possibility Claim is true.

What should we say of such a case? Let’s introduce the idea of a *subjectively idealised conative state.* These are relevantly like subjectively idealised desires, except they apply to other sorts of conative states such as hopes, fears, anticipations and the like. Then we are inclined to say that P1’s subjectively idealised anticipatory states (one kind of subjectively idealised conative state) are such that P1 will ideally anticipate *all* the experiences/desires of its adjacent agent-stage. To be clear, by this we mean that P1 will anticipate having *whatever* P2’s experiences are and having *whichever* desires P2 has). Why think so?

Suppose that P1 desires and intends to sit down, and that it is P2 who sits. As part of desiring and intending to sit, P1 anticipates sitting. So, P1 anticipates *some* of P2’s experiences. Now suppose that while P2 sits, as a matter of fact P2 spills some hot coffee on its leg. Suppose that we inform P1 of this. Will P1 come to anticipate being (mildly) burned? Surely it will. Once we idealise P1’s relevant conative states, they will include the anticipation of mild coffee-induced burning. What of other experiences of P2’s? Suppose that we inform P1 that the while sitting down, P2 is hungry. Will P1 anticipate being hungry while sitting down? Surely it will. Likewise for any other experience of P2’s that we can list.

Now suppose we tell P1 that as P2 is sitting down, it desires a muffin. Will P1 come to anticipate having this desire? Surely the answer is yes. So, we say, the following is true. Necessarily, if an agent-stage, P, anticipates some of the experiences/desires of a continuer P\*, then P’s subjectively idealised conative states will be such that P anticipates *all* the experiences/desires of that continuer. Since we’ve already argued that, necessarily, an agent-stage P anticipates some of the experiences/desires of its adjacent agent-stage; it follows that, necessarily, an agent-stage has subjectively idealised conative states of anticipating all the experiences/desires of its adjacent continuer. Given this, it follows that if P1 has reason to promote the wellbeing of adjacent agent-stage P2, then it also has reason to promote the wellbeing of all its non-adjacent continuers. Since we’ve argued that P1 does have reason to promote the wellbeing of its adjacent agent-stage, it follows that it has reason to promote the wellbeing of all its non-adjacent continuers.

Here is our second argument for the falsity of the Non-Adjacent-Stage Possibility Claim. Again, we need to back off from our claim that we will remain neutral regarding which theory of wellbeing is the right one. Suppose, again, we were to adopt a desire satisfaction theory of wellbeing according to which wellbeing consists in the satisfaction of desires.[[38]](#footnote-39) And suppose, again, that one thinks that the person-stage who benefits from a desire being satisfied is the person-stage that exists at the time-of-satisfaction. So if it is P1 who desires to sit, and P2 who satisifes this desire, it is P1 whose wellbeing is promoted. If one endorses this view, then adjacent agent-stages such as P2 promote the wellbeing of earlier stages, such as P1, when they satisfy their desires. This means that P1s wellbeing in part depends on the wellbeing of P2, since it is P2 who can promote P1’s wellbeing. Indeed, P3’s wellbeing depends on that of P4, and P2’s on P3, and P1’s on P2. So, if there is a chain of adjacent continuers, then any stage in that chain has reason to promote the wellbeing of *any* continuer in that chain, because any agent-stage has a reason to promote the wellbeing of its adjacent continuer and promoting the wellbeing of later continuers is a way to promote the wellbeing of adjacent continuers, and so there is a chain of reasons that ‘percolate backwards’ from later to earlier stages.

So, insofar as theories of wellbeing are, if true, necessarily true, it follows that necessarily, for any agent-stage P such that P\* is P’s continuer and P\* does not overlap P’s adjacent continuer, P’s actual/idealised desires provide P with a reason to promote the wellbeing of P\*. The Non-Adjacent-Stage Possibility Claim is false.

Again, we do not claim that one cannot resist these arguments. Clearly one can: one can deny that the right account of wellbeing is desire satisfaction; and one can deny that it is the stage whose desires are satisfied whose wellbeing is promoted; and one can deny that, necessarily, stages will have subjectively idealised anticipations of all the desires/experiences of their adjacent continuers. With regard to this last claim the perdurantist might, again, retreat to a weaker version of the axiom and argue that persons that are at all psychologically like us do have these patterns of anticipations (or at least, ideally do) and that this means that any person remotely like us has reason to promote the wellbeing of all its non-adjacent continuers.

**7. Overall Wellbeing**

Recall that the strong perdurantist axiom of prudence says the following:

**Strong Perdurantist Axiom of Prudence**: necessarily any agent-stage is rationally required to promote the *overall* wellbeing of each of its continuers.

Nothing we have said so far vindicates this axiom. Indeed, we are sceptical that anything as strong as this axiom can be vindicated by appealing only to instrumental reasons. Still, perhaps the perdurantist can vindicate something in the rough vicinity of this axiom. And, you might think that capturing something close to this axiom is important, because the account we have so far offered is not really faithful to how we think about our reasons to promote wellbeing.

To see this, consider agent-stage P. Suppose that P desires to sit down, and that it is P\* who will satisfy that desire. Then, we say P has a reason to promote P\*’s wellbeing. But of course, it’s entirely open that P’s desires might make it rational for P to retard P\*’s wellbeing in a host of non-sitting-down relevant ways. Yet this will no doubt seem wrong to most of us. Moreover, it does not seem true to our lived experience. We don’t simply promote the wellbeing of our continuer stages because they satisfy our desires, and we don’t take ourselves to have reason to promote their wellbeing *only* insofar as they satisfy our desires.

We think that what is wrong with the picture above is not that we all have, and take ourselves to have, reasons to promote the wellbeing of our continuers that do not issue from our desires, but rather, that the picture above presents a tremendously impoverished view of our desires.

There are multiple kinds of desires that agent-stages can have. An agent-stage can have impersonal desires.[[39]](#footnote-40) I impersonally desire world peace if I care that there will be world peace, and I don’t care whether the agent-stages who bring it about are my continuers or someone else’s. Sometimes, as Perry (1976) notes, as a matter of empirical fact the agent-stages best placed to satisfy these kinds of desires are one’s own continuers. So perhaps I desire that a certain project come to fruition, a project that I care about, and have begun already and spent many years on. If I impersonally desire that the project be completed, then I don’t care that I am the one to complete it. Still, given the way the world is, it is much more likely that the project will be completed if my continuers complete it. And this, says Perry, gives me reason to promote the wellbeing of those continuers. Perry attempts to marshal this idea to explain why we have reason to promote the wellbeing of our continuers. He’s prepared to concede that insofar as our projects can be completed without our promoting the wellbeing of certain future continuers, then we have no reason to promote the wellbeing of those continuers.

Importantly though, many of our desires are not impersonal in this manner. I desire that Annie (a labradoodle) be walked. But I also desire that I, myself, take Annie for a walk. Parents desire that their children are cared for; most of them desire that they are the ones doing the caring. We will call such desires *personal desires*. While some desires can be either personal or impersonal, such as the desire that one’s children are taken care of, some desires seem to be necessarily personal. The desire for food is not the desire that someone, somewhere, eat. It’s a desire that one has continuers that eat. Many of our desires are like this: they can only be satisfied by our continuers.

This is important, since it often seems, from a first-personal perspective, as though we have reason to promote the wellbeing of our continuers even if they will not be best placed to satisfy our impersonal desires. And that, of course, is right: it is right because it is only they that can satisfy our personal desires.

That, however, cannot be the full story, for it often seems to us as though we have reason to promote the wellbeing of (at least adjacent) continuers, *regardless of whether they satisfy our desires.* This is the sense in which you might think that we take ourselves to have reason to promote the *overall* wellbeing of our continuers rather than simply reason to promote their wellbeing insofar as doing so satisfies our own desires.

We think that there is something clearly right about this. But we think that this is because many, and perhaps most, of our desires are in fact personal ones and because many of those personal desires, motivated by anticipatory states, are desires that continuers have certain sorts of experiences and lack others. As a matter of fact, we want our continuers to have pleasant experiences (*inter alia*) precisely because we anticipate those experiences.

Indeed, if we are right and, necessarily, agent-stages ideally anticipate all of the experiences/desires of their adjacent stages, then it’s hardly surprising that we aim to promote something *like* their overall wellbeing. At the very least we desire that things go well for our continuer stages, in ways that outstrip merely caring that they go well so that those stages can satisfy what we might think of as our ‘local’ or ‘specific’ desires—desires such as eating, or sitting down, or finishing a paper. Rather, we have much more general, personal desires that things go well for our continuers, and this gives us reason to promote something that looks closer to their overall wellbeing.

Given this, something a bit like the strong axiom is true. It’s *not* that all agent-stages have reason to promote the overall wellbeing of their continuer stages. But agent-stages that are at all psychologically like us in anticipating the experiences of continuer stages will have instrumental reason to promote the wellbeing of all of their continuer stages in a much fuller and more general sense. And, we think, that may well be enough to capture what is attractive about the strong axiom.

**8. Conclusion**

We have argued that any agent-stage has instrumental reason to promote the wellbeing of all its continuers. Hence, purely instrumental rationality can ground perdurantist prudence. Perhaps there are also *sui generis* prudential reasons, and the perdurantist can avail herself of these. But, we say, it does not matter if there are not, and she cannot. For appeal to instrumental reasons will suffice. Hence perdurantism cannot be dismissed on the grounds that it cannot account for prudential rationality.

**References**

Arntzenius, F. 2008. No Regrets, or: Edith Piaf Revamps Decision Theory. *Erkenntnis* 68(2): 277-297

Barry, B. 1965. *Political Argument*. London: Routledge & Kegan Paul.

Beck, S. 2015. The Extreme Claim, Psychological Continuity and the Person Life View. *South African Journal of Philosophy* 34 (3):314-322.

Brink, David O. 1997. Rational Egoism and the Separateness of Persons. *Reading*

*Parfit* , ed. Jonathan Dancy. Oxford: Blackwell: 96–134.

Braddon-Mitchell, D. and Miller, K. 2004. How to be a Conventional Person. *The Monist* 87(4): 457-474.

—— 2020a. Conativism about personal identity. In A. Sauchelli (Ed.), *Derek Parfit's Reasons and Persons: An introduction and critical inquiry*, 129-159. London: Routledge.

—— 2020b. Surviving, to some degree. *Philosophical Studies*, 177, 3805-3831.

Braddon-Mitchell D. and West, C. 2001. [Temporal Phase Pluralism.](https://philpapers.org/rec/WESTPP) *Philosophy and Phenomenological Research* 62(1): 59–83.

Brandt, R. B. 1966. The Concept of Welfare. In *The Structure of Economic Science*, ed. S. R. Krupp (ed.), Englewood Cliffs, NJ: Prentice-Hall, 257–76.

Butler, Joseph. 1975. Of Personal Identity. In *Personal* Identity, ed. J. Perry, 99–105. First published in 1736.

Carter, Ian. 2018. Equal Opportunity, Responsibility, and Personal Identity. *Ethical Theory and Moral Practice* 21: 825–839.

Dietz, A. 2020. Are My Temporal Parts Agents? *Philosophy and Phenomenological Research* 100 (2):362-379.

Dorsey, D. 2018. Prudence and Past Selves. *Philosophical Studies* 175(8): 1901–1925.

Dougherty, T. 2011. On Whether to Prefer Pain to Pass. *Ethics*. 121(3): 521-537.

—— 2015. Future-Bias and Practical Reason. *Philosophers’ Imprint* 15/30: 1–16.

Enoch, D. 2006. Agency, Schmagency: Why Normativity Won’t Come From What Is Constitutive of Action. *Philosophical Review* 115 (2):169-198.

Fletcher, G. 2021. *Dear Prudence: The Nature and Normtivity of Prudential Discourse.* Oxford University Press.

Greene, P. and M. Sullivan. 2015. Against Time Bias. *Ethics* 125/5: 947–70.

Hare, C. 2007. Self-Bias, Time-Bias, and the Metaphysics of the Self and Time. *Journal of Philosophy* 104: 350–73.

—— 2008. A Puzzle about Other-Directed Time-Bias. *Australasian Journal of Philosophy* 86: 269–77.

—— 2013. Time – The Emotional Asymmetry. In *A Companion to the Philosophy of Time*, ed. A. Bardon and H. Dyke, Wiley Blackwell: 507–20.

Hare, R. M. 1981. *Moral Thinking*. Oxford: Clarendon Press.

Harman, E. 2009. I’ll be glad I did it. Reasoning and the Significance of Future Desires.” *Philosophical Perspectives* 23(1):177-199

Harsanyi, J. 1977. Morality and the Theory of Rational Behavior. *Social Research* 44: 623– 56.

Hedden, Brian. 2015. *Reasons without Persons: Rationality, Identity, and Time*. Oxford:

Oxford University Press.

Hurley, Susan. 1989. *Natural Reasons*. New York: Oxford University Press.

Jackson, Frank. 1987. Group Morality. In *Metaphysics and Morality*, ed. Philip

Pettit, Richard Sylvan, and Jean Norman. Oxford: Blackwell. 91–110.

Johnston, M. 2017. The Personite Problem: Should Practical Reason be Tabled? *Noûs* 50: 617-644.

—— 1989. “Relativism and the Self”.  In [Relativism: Interpretation and Confrontation](https://philpapers.org/rec/KRARIA-3), ed., M. Krausz, Notre Dame University Press.

—— 2010. *Surviving Death*. Princeton, N.J.: Princeton University Press.

Kaiserman, A. 2019. Stage Theory and the Personite Problem. *Analysis* 79: 215-222

Kleinig, John. 1983. *Paternalism*. New Jersey: Rowman & Allanheld.

—— 2009. *Paternalism and Person Identity*. Jahrbuch für Wissenschaft und Ethik 14: 93–106.

Kogan, Terry S. 1976. The Limits of State Intervention: Personal Identity and Ultra-Risky Actions. *Yale Law Journal* 85: 826–846.

Kovacs, D. 2016. Self-Made people. *Mind* 125: 1071-1099

—— 2020. Diachronic Self Making. *Australasian Journal of Philosophy* 98:349-362.

—— (forthcoming). “Self making and subpeople.”

Lin, E. 2015. Prudence, Morality and the Humean Theory of Reasons. *Philosophical Quarterly* 65: 220-40

Longenecker, M. 2020a. Perdurantism, Fecklessness and the Veil of Ignorance. *Philosophical Studies* 177: 2565-2576.

——2020b. Is Consequentialist Perdurantism in Moral Trouble.? *Synthese* 198 (11):10979-10990.

Longenecker, M. (forthcoming). Community-Made Selves. *Australasian Journal of Philosophy*.

Markovits, J. 2014. *Moral Reason*. Oxford University Press.

Martin, R. 2008. *Self-Concern: An Experiential Approach to What Matters in Survival.* New York: Cambridge University Press.

McCall, S. and Lowe, D. J. 2002. 3D/4D Equivalence, the Twins Paradox and Absolute Time. *Analysis* 63:114–123.

Merricks, T. 1999. Endurance, Psychological Continuity, and the Importance of Personal Identity. *Philosophy and Phenomenological Research* 59: 983-997.

Miller, K. 2007. *Issues in theoretical diversity: persistence composition and time.* Springer.

—— 2009a. Deciphering Personal-identity Conventionalism. Special Issue *Rivista di Estetica “Convenzioni”*,eds. Elena Casetta and Achille Varzi, 44: 59-85.

——2009b). Ought a four-dimensionalist to believe in temporal parts? *Canadian Journal of Philosophy* 39: 619-646

—— 2010. Persons as *Sui Generis* Ontological Kinds: Advice to Exceptionists. *Philosophy and Phenomenological Research* 81: 567-593.

——2013. Personal identity minus the persons. *Philosophical Studies*. 166: 91-109.

Nozick, Robert. 1981. *Philosophical Explanations*, Cambridge, MA: Harvard University

Press.

Olson, Eric T. 2002. Thinking Animals and the Reference of ‘ I’. *Philosophical Topics*, 30: 189– 207.

Olson, E. 2007. *What Are We? A Study in Personal Ontology*. Oxford: Oxford University

Press.

——Olson, E. 2010. Ethics and the generous ontology. *Theoretical Medicine and Bioethics* , 31: 259– 270.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

Pearson, O. 2018. Appropriate emotions and the metaphysics of time. *Philosophical Studies*, *175*: 1945-1961.

Perry, J. 1976. The importance of being identical. In *The identities of persons,* ed., A. O. Rorty. Berkeley: University of California Press.

Prior, A.N. 1959. Thank Goodness That’s Over. *Philosophy* 34: 12–17.

Ramsey, F. P. 1926. Truth and Probability. In *Foundations of*

*Mathematics and Other Essays*, ed. R. B. Braithwaite. London: Kegan, Paul, Trench, Trubner, & Co., 1931, 156–198.

Rawls, J. 1971. *A Theory of Justice*. Cambridge, MA: Belknap

Reagan, Donald H. 1983. “Paternalism, Freedom, Identity, and Commitment”. In *Paternalism*, ed. Rolf E. Sartorius, 113–138. Minneapolis: University of Minnesota Press.

Rovane, Carol. 1998. *The Bounds of Agency: An Essay in Revisionary Metaphysics .*

Princeton, N. J.: Princeton University Press, 1998.

Schechtman, M. 1996. *The constitution of selves*. Ithaca, N.Y.: Cornell University Press.

Schechtman, M. 2005. Experience, Agency, and Personal Identity. *Social Philosophy and Policy* 22:1-24.

Schechtman, M. 2014. *Staying Alive: Personal Identity, Practical Concerns and the Unity of a Life.* Oxford University Press.

Schroeder, M. 2007. *Slaves of the Passions*. Oxford University Press.

Sebo, Jeff. 2015a. The Just Soul. *Journal of Value Inquiry*, 49: 131– 143.

—— 2015b. Multiplicity, self-narrative, and akrasia. *Philosophical Psychology*, 28:

589–605.

Shaw, W. 1999. *Contemporary Ethics*. Malden, MA: Blackwell Publishers.

Shoemaker, David. 1999. Selves and Moral Units. *Pacific Philosophical Quarterly* 80:

391–419.

Sidgwick, H. 1884. *The Methods of Ethics*. London: Macmillan.

Singer, P. 1979. *Practical Ethics*. Cambridge: Cambridge University Press.

Slors, M. 2004. Care for One’s Own Future Experiences. *Philosophical Explorations* 7: 183-195.

Slote, M. 1989. *Beyond Optimizing. A Study of Rational Choice*. Harvard University Press.

Smith, A. 1790. *The Theory of Moral Sentiments*. Oxford: Oxford University Press.

Sullivan, M. 2018. *Time Biases: a Theory of Rational Planning and Personal Persistence*. Oxford University Press.

Taylor, A. P. 2020. The painful implications of four-dimensionalism. E*thics, medicine and public health*. 13: 1-10.

Taylor, A. P. 2013. The frustrating problem for four-dimensionalism. *Philosophical Studies* 165: 1097-1115.

Unger, P. 1990. *Identity, Consciousness and Value.* Oxford University Press.

Van Frassen, Bas. 1984. Belief and the Will. *The Journal of Philosophy* 81: 235–256.

Von Wright, G. H. 1963 *The Varieties of Goodness*. New York: The Humanities Press.

West, C. 1996. *Surviving Matters*. PhD Dissertation, ANU.

White, S. 1989. Metapsychological Relativism and the Self. *Journal of Philosophy* 86: 298-325.

Whiting, J. 1986. Friends and Future Selves. *Philosophical Review* 95: 547-80.

Williams, B. 1981. Internal and External Reasons. In *Moral Luck,* Oxford University Press, 101-113.

Zimmerman, D. 2012. Personal identity and the survival of death. In *The Oxford handbook of philosophy of death*, eds. B. Bradley, F. Feldman, J. Johansson, 97-154, Oxford University Press.

1. For defences of perdurantism see Quine (1950), Lewis (1976) and (1986), Armstrong (1980), Heller (1984) and (1990), Hudson (2001), Sider (2001), Miller (2007), Wasserman (2016), Balashov (2000) and Hales & Johnson (2003). There are other, related, views on which persons are four-dimensional but lack temporal parts (for discussion see Miller (2009b)), but we set those aside. [↑](#footnote-ref-2)
2. For responses to these kinds of worries see Longenecker (2020a) and (2020b) and Kaiserman (2019). [↑](#footnote-ref-3)
3. Exdurantism or stage-theory is the view that persons just are stages, and claims about how they were, or will be, are made true by the existence of temporal counterparts. So on such views persons are the fundamental seat of agency. But, arguably at least, arguments concerning subpeople do not arise in quite the same way within a stage-theoretic picture. In this paper we take persons to be four-dimensional entities. But everything we say about how to think about prudence within a perdurantist framework can be adopted very naturally by stage-theorists. [↑](#footnote-ref-4)
4. We do not suggest that this, alone, is sufficient to resolve the problem of subpeople, and that is not the aim of this paper; but we do think it’s an important first step. [↑](#footnote-ref-5)
5. See for instance Sidgwick (1884). [↑](#footnote-ref-6)
6. Slote (1989). [↑](#footnote-ref-7)
7. Parfit (1984). [↑](#footnote-ref-8)
8. While most philosophers hold that near-bias (that is, more highly valuing the wellbeing of near selves over far ones simply on the basis of where those selves are located in time) is rationally impermissible (see for instance Smith (1790), Rawls (1971), and Sidgwick (1884, 380-381)), many hold that it is rationally permissible (or even obligatory) to more highly value the wellbeing of future selves over past selves (see for instance Prior (1959), Hare (2007) and (2013), and Pearson (2018))—though also see Dougherty (2011) and (2015), Greene and Sullivan (2015), Sullivan (2018), and Brink (2011) for views to the contrary. [↑](#footnote-ref-9)
9. That is, at least some of P\*’s parts are later than any of P’s parts. [↑](#footnote-ref-10)
10. See for instance Dorsey (2018). [↑](#footnote-ref-11)
11. Notice this is distinct from the question of whether one maximises (or not) overall or partial wellbeing. [↑](#footnote-ref-12)
12. More carefully, the ultimate grounds will be P’s non-instrumental desires. [↑](#footnote-ref-13)
13. For defences of Humeanism about instrumental reason see Schroeder (2007), Markovits (2014), Williams (1981), and Lin (2015). [↑](#footnote-ref-14)
14. The view we have in mind is one on which not only can the idealisation process change which instrumental desires one has, as one might think is occurring in the case of the canine companion (where I non-instrumetnally desire to be, say, happy, and a canine companion would satisfy this desire); but it can also change one’s non-instrumental desires. That is, under conditions I take to be better, I would have diffeerent non-instrumental desires. [↑](#footnote-ref-15)
15. For defences of this view see Parfit (1984), Sullivan (2018), and Fletcher (2021). [↑](#footnote-ref-16)
16. See for instance Parfit (1984) and Fletcher (2021). [↑](#footnote-ref-17)
17. Though in fact Sidwick (1884, 418) thinks a similar problem arises even if persons endure. [↑](#footnote-ref-18)
18. Sentiments such as this go back at least as far as Butler (1784). [↑](#footnote-ref-19)
19. Rather ironically, the idea that persons perdure has been used to argue that liberals can legitimately intervene in cases otherwise conceived of as self-harm (and hence falling under the banner of ‘autonomy’) precisely because the relelvant future selves that bear the harm are numerically distinct from the self perpetuating the harm. See for instance Kleinig (1983) and (2009), Reagan (1983), Kogan (1976), and Carter (2018). [↑](#footnote-ref-20)
20. The claim that persons do not have these kinds of *sui generis* prudential reasons to promote the welfare of other person-stages is often known as the extreme claim. See Parfit (1984, 307) for discussion. That claim is often marshalled not against perdurantist views of persistence, but more broadly against reductionist views of personal-identity (views on which personal identity reduces to the obtaining of certain mental and physical properties of person-stages, and the relations between said stages). See for instance Parfit (1984); Schechtman (1996), (2005), (2013); and Butler (1784). The dialectic here is a little murky. Sometimes these kinds of objections to reductionist views seem to simply suppose that reductionists must in fact be committed to something like perdurantism, on the grounds that the continuity relations that reductionists appeal to clearly are not identity relations. Certainly, there are those who argue that endurantists cannot endorse reductionism precisely because identity does not reduce to, or consist in, the obtaining of these other relations (see for instance Merricks (1999)). On the other hand, some reductionists clearly do seem to think that their view is consistent with strict identity over time (at least as long as the account includes a non-branching rquirement), and take themselves to be giving an account of what that identity consists in (see for instance Parfit (1984)). But if numerical identity over time in some sense consists in the obtaining of these kinds of continuity relations, then it’s unclear why reductionists have any more of a problem than non-reductionists in explaining what grounds prudential rationality. At any rate, we are interested in this argument as it pertains to perdurantism, rather than as it pertains to reductionist views in general. [↑](#footnote-ref-21)
21. Something *like* this argument has been offered by, *inter alia*, Schechtman (1996), Butler (1784), and Wolf (1986, 719); and perhaps underlies some of Korsgaard’s comments in Kordsgaard (1989, 113-114). It has motivated some to suggest that while in general objects perdure, persons are special and persist in some other manner. See, for instance, Miller (2010). [↑](#footnote-ref-22)
22. Indeed, whether numerical identity over time really is sufficient to ground *sui generis* prudential reasons is open to debate. Whiting (1986), Beck (2013), Sidgwick (1884, 418), and Miller (forthcoming) do not think so. [↑](#footnote-ref-23)
23. Where time and personal time come apart, continuers and adjacent continuers (etc.) shoud be defined in terms of personal-time not time itself. We leave these complications aside. [↑](#footnote-ref-24)
24. Views of this kind have been defended by Kovacs (2016), (2020); Zimmerman (2012); Johnston (2010, 5); and Whiting (1986). They are sometimes known as conventionalist (see Kovacs (2016) and (2020), Braddon-Mitchell and Miller (2004), Miller (2009a), Longenecker (forthcoming), and Schechtman (2014)) or conativist (see Braddon-Mitchell and Miller (2020a), (2020b)) or practice-dependent (see Braddon-Mitchell and West (2001), and West (1996)). [↑](#footnote-ref-25)
25. See for instance Braddon-Mitchell and West (2001), Braddon-Mitchell and Miller (2020a) and (2020b), and Miller (2009a). [↑](#footnote-ref-26)
26. For discussion of desires such as these see Brink (1997), Perry (1976), and Hurley (1989). [↑](#footnote-ref-27)
27. Though see Duncan and Miller (2015) who defend the view that mere similarity might do the trick. [↑](#footnote-ref-28)
28. Thanks to Antony Eagle for this suggestion. [↑](#footnote-ref-29)
29. See Van Fraassen (1984). [↑](#footnote-ref-30)
30. See Arntzenius (2008). [↑](#footnote-ref-31)
31. Following Harman (2009) an even weaker version of reflection for desires which appeals to what one *reasonably* believes, and to what one’s future self’s *reasonable* desires are, yields the following:

**Reflection for Desires**: If P reasonably believes that P’s future self will reasonably prefer that x be true, and if P reasonably believes that P’s future self will not be in a worse epistemic or evaluative position at that time, then P should now prefer that x be true. [↑](#footnote-ref-32)
32. Or, if you prefer, reasonbly anticipates reasonably desiring. [↑](#footnote-ref-33)
33. See for instance Harman (2009) who denies it. [↑](#footnote-ref-34)
34. If you think there are possible agent-stages whose even idealised desires will not be such that they care about lost evidence, or irrationality, then those aspects of the principle could be jettisoned for our purposes. [↑](#footnote-ref-35)
35. See Harsanyi (1977), von Wright (1963), Barry (1965), Brandt (1966), Ramsey (1926), Singer (1979), and Hare (1981) [↑](#footnote-ref-36)
36. Shaw (1999, 53) [↑](#footnote-ref-37)
37. Dorsey (2013) endorses this view. [↑](#footnote-ref-38)
38. See Harsanyi (1977), von Wright (1963), Barry (1965), Brandt (1966), Ramsey (1926), Singer (1979), and Hare (1981). [↑](#footnote-ref-39)
39. Following the terminology of Whiting (1986). [↑](#footnote-ref-40)