What Justifies Our Bias Toward the Future?*

Todd Karhu

ABSTRACT. A person is biased toward the future when she prefers, other things being equal, bad events to be in her past rather than her future or good ones to be in her future rather than her past. In this paper, I explain why both critics and defenders of future bias have failed to consider the best version of the view. I distinguish external time from personal time, and show that future bias is best construed in terms of the latter. This conception of future bias avoids several standard objections. I then consider a new justification of future bias which is consistent with that construal. My discussion points to a new position regarding the basic relation that grounds rational egoistic concern over time, according to which that relation is asymmetric between person-stages. I also explain how this way of justifying future bias would resolve the apparent tension between the future bias we display in our own case and our relative indifference to the timing of the good and bad things that happen to other people.

1. Introduction

Most of us are familiar with the impulse to postpone bad things for as long as possible and obtain good things as soon as we can. Indeed, that inclination is often present even when the good things would be better if waited for and delaying the bad things will only make them worse. This *bias toward the near* is something many people expend considerable effort fighting against. For we seem to think that *ideally* we would be neutral between the near and further future.

But human beings display a second psychological bias concerning the timing of good and bad things, no less pronounced but much less frequently confronted. This is the preference for good experiences to be in the future rather than the past, and bad ones to

* This is a preprint of an article whose final and definitive form will be published in *the Australasian Journal of Philosophy* [2022]. *The Australasian Journal of Philosophy* is available online at: http://www.tandf.co.uk/journals/.

be in the past rather than the future. Here is a classic illustration of the phenomenon:

> *Your Past or Future Operation*: You have been admitted to the hospital for a painful but otherwise safe operation. So painful is this operation, in fact, that patients are afterwards given a drug which causes them to forget the entire thing. Now you wake up, unable to recall whether you have had the operation yet. The nurse you hail can tell you only this. If it was yesterday, it lasted for ten hours, and if it is tomorrow, it will take one hour [Parfit 1984: 165–6].

Upon finding themselves in this predicament, many people would hope to have already had the operation, even though it will then have been ten times as long (see [Greene et al. 2021b]). This is an instance of a more general phenomenon. When it comes to negative experiences like pain, we are *biased toward the future*: we want them to be in the past rather than the future, even when that means they will have been longer or more unpleasant [Caruso et al. 2008; Lee et al. 2020; Latham et al. 2021; Greene et al. 2021a, 2021b]. For positive experiences, this pattern is reversed. We want our pleasures to be ahead of us rather than behind us, even when that means they will be less pleasurable for it [Caruso et al. 2008; Lee et al. 2020; Greene et al. 2021a, 2021b, forthcoming].

Unlike the bias toward the near, future bias will probably strike most people as perfectly sensible. But many philosophers doubt that it is a justifiable pattern of preferences to have. In this paper, I will explain how some prominent arguments for and against future bias are based on a mistaken understanding of our actual attitudes toward the timing of the good and bad events in our lives. I will then explain how the kind of future bias we in fact display, which involves a special concern for what happens in the personal future, can be justified as part of a general notion of what it takes for something to be in or against someone's self-interest *per se*. I will also show how this way of justifying future bias reconciles it with our apparent indifference about whether the good and bad things that happen to other people are in the past or the future.[1]

---

[1] There is one kind of argument against future bias that I lack the space to consider here. Pragmatic arguments against future bias aim to establish that future bias is practically irrational because there are conceivable situations in which a future-biased agent will make a sequence of choices which leaves her worse off whatever happens—provided she exhibits some other ostensibly rational disposition like risk aversion. The key pragmatic arguments are Dougherty [2011], Greene and Sullivan [2015], and Sullivan [2018]. For defences of future bias against those arguments, see Hare [2013], Dorsey [2017], Tarsney [2017], and Scheffler [2021].

2. The Metaphysics of Time and the Bias Toward the Personal Future

In seeking to defend future bias, it is natural to look to metaphysical differences between the past and the future that might justify treating events in one differently from events in the other. One conspicuous difference is that unlike the past, we can typically alter the future. This seems unlikely to be a viable defence of our actual attitudes, however, for it seems we would not ordinarily care much less about future pain when and because we knew it to be inevitable [Parfit 1984: 168–70; Garrett 1988: 204; Kamm 1993: 28; Hare 2013: 512; Dorsey 2018: 1910]. Another difference is that the future may be metaphysically open, in the sense that contingent facts about how things will be are not presently settled. But this, too, seems unviable: if anything, it seems we should want *bad* things to be unsettled and *good* ones to be settled [Hare 2013: 511; Yehezkel 2014: 77].[2]

Proponents of future bias often focus on what is arguably an even more basic difference between the past and the future. According to the A-Theory of time, whether an event is past, present, or future is an intrinsic, observer-independent property of it. Time objectively passes, which involves future events becoming present, present events becoming past, and past events becoming ever more so. It is widely thought that belief in the A-Theory both explains, and if true justifies, the widespread bias toward the future. If time objectively passes, then unlike past goods and bads, future ones will eventually be present. Since these things matter precisely because of what it is like to have them in the present, it makes perfect sense that if the A-Theory is true we should want the bad things in life to be in the past and the good ones in the future [Prior 1959; Schlesinger 1976; Parfit 1984: 178–9; Kamm 1993: 28; Hare 2007, 2013].

Notwithstanding its intuitive appeal, however, this justification of future bias faces three difficulties. The first is just that the A-Theory appears unable to function as an independent justification for future bias. After all, what is it to say that pain matters more when and because it will be present, if not just that it matters more when and because it is in the future? For this reason, those attracted to this argument usually characterise it less as a justification of future bias than as a way of emphasising that if the A-Theory is true it is beyond justification: it is a brute fact of prudential rationality that things that will be present should *ipso facto* matter more to us than things that will not be [Parfit 1984: 178, 181; Heathwood 2008: 62; Hare 2013: 514–15; Dorsey 2017: 352–3].

---

[2] Hare's article provides an instructive overview of several other metaphysical justifications for future bias, which he argues ultimately prove untenable.

The second difficulty with the argument from the A-Theory is that it depends on there actually *being* some intrinsic property of future events that differentiates them from past ones. Yet according to the competing B-Theory, despite all appearances to the contrary, there is no observer-independent fact of the matter as to whether an event is in the past or the future. While events can be past or future relative to others—which is to say, earlier or later than them—these properties are relational only, not ones belonging to the events themselves.

The B-Theory is often said to pose a serious threat to future bias wholesale [Parfit 1984: 177–81; Prior 1959; Schlesinger 1976; Garrett 1988: 204; Cockburn 1998; Hare 2007: 360–1, 2013: 513–15; Greene and Sullivan 2015: 953; Sullivan 2018: 108–9]. For one thing, if belief in the A-Theory explains our future biased attitudes, there would seem little reason to expect them to be independently justified if they turn out to be based on a metaphysical view that is false. More critically, however, many authors simply find it hard to see how an event's being in the future could be in itself normatively significant if there is no intrinsic property of future events that differentiates them from past ones [Parfit 1984: 179; Prior 1959; Schlesinger 1976; Garrett 1988: 204; Cockburn 1998; Yehezkel 2014: 74; Greene and Sullivan 2015: 953]. On a B-Theoretic picture of reality, events are spread out in time much like locations are in space, a point often emphasised by analogising 'now' to 'here'. There is thus a real sense in which the painful surgery I had in 2018 is still taking place—just not at the time at which I happen to be pondering it. If that is so, then as Preston Greene and Meghan Sullivan put it, 'distinguishing between past and future experiences can seem just as arbitrary as distinguishing between experiences that happen here and experiences that happen there' [2015: 953].

There is a third problem with justifying future bias on the basis of the A-Theory: the argument misattributes to human beings a particular kind of future bias.[3] Call it the bias toward the *external future*. A person is biased toward the external future at a given time if she prefers, other things being equal, that the good things which happen to her occur later than that time and the bad ones occur earlier than it. But a person can be future-biased in another way. Rather than caring especially about what happens in *the* future, a person might care especially about what happens in *her* future. She might want her pains to occur earlier *in her life* than the present moment,

[3] I initially argued we are not biased toward the external future in a PhD thesis [Karhu 2019: 51–57], but a similar argument has been developed independently by Kristie Miller [2021]. Miller's article contains many important nuances that I do not cover here. Unlike me, Miller also expresses scepticism that we are biased toward the personal future—a difference I discuss later on.

and her pleasures to occur later in it. Call this the bias toward the *personal future*.[4]

Now, in all actual cases of which we are aware, an individual's personal future coincides with the external future. In middle age, one's old age is later in external time and one's adolescence earlier in it. But when, in imagined cases, the two measures of time come apart, in fact it is an event's location in personal time, not external time, that affects our prudential concern about it. If that is right, then truths about the metaphysics of time neither support nor challenge the kind of future bias worth defending, because that kind has nothing to do with the structure of external time.

Consider the following pair of cases:

> *Business as Usual*: Tomorrow, you must undergo ten hours of excruciating pain.
>
> *To the Middle Ages*: Tomorrow, you will travel back in time to the year 1350, whereupon your arrival you must undergo ten hours of excruciating pain.

Some think that deep philosophical or physical investigation reveals time-travel to be impossible, but even if so, we can ask what our pre-theoretic attitudes about these cases are. And if you are like me, your attitude toward the ten hours of pain in the second case is like your attitude in the first. Hence, we have an episode of pain that is earlier in external time than the present, but which seems no more desirable than an identical episode that is later than it.

Now take another pair of cases:

> *More Business as Usual*: Some time in your youth, you underwent ten hours of excruciating pain.
>
> *Already in the Middle Ages*: Some time in your youth, you underwent ten hours of excruciating pain. You have since travelled back in time to the fourteenth century.

Consider *Already in the Middle Ages*. Imagine that you are now in the fourteenth century and are pondering the ten hours of excruciating pain in your youth. If you are like me, the fact that this pain is in the external future does very little to make you care more about it. So we have a pair of cases that seems to show that pain that is later in external time matters little more to us than identical pain that is

---

[4] The distinction between external time and personal time comes from David Lewis [1976].

earlier in it. Admittedly, there may seem a certain fatalism to the pain *Already in the Middle Ages* (one that was not present in *To the Middle Ages*). Even though it is behind you in personal time, in a way it will come again. But even that sense of discomfort, I suspect, is very minor compared to your feelings toward the painful episode itself.

Most discussants of future bias have taken it to be a pattern of preferences about where good and bad things are located in external time. On the contrary, however, assuming the preceding reactions would be widely shared, we are more-or-less indifferent to when in external time these things occur. Accordingly, whether there are deep metaphysical differences between the past and the future is not relevant the justifiability of future bias. Appeals to the A-Theory are beside the point, and since our actual attitudes do not track the structure of external time, those attitudes are equally compatible with the B-Theory and its negation [Karhu 2019].

The best explanation for our reactions to these cases, I believe, is that what really matters to us is whether the pain happens in our personal future or personal past. Kristie Miller [2021] disagrees. Although she shares the prediction about how most people will react to a case like *To the Middle Ages*, she suspects that if people believed their past pain could be retroactively affected, they would be no less concerned about it than they would be about a similar pain in their personal future. To show this, Miller first asks the reader to imagine deciding whether or not to have 1 unit of pain tomorrow. If you do choose to have it, a trustworthy party will change the past so that you had 5 units of pain yesterday, making 6 units in total. If, alternatively, you opt for no pain tomorrow, the party will instead make it the case that you had 7 units of pain yesterday. Assuming that you temporarily cannot recall yesterday's pain, so as to factor out the possibility that your desire to avoid having a more painful memory will influence your decision, Miller tentatively predicts that you will select the 1 unit of pain tomorrow, thereby opting for 6 rather than 7 units in total.

For what it is worth, I can say that my own intuition diverges sharply from hers. If I had an hour of intense pain ahead of me, to avoid it I would readily make it the case that I had 7 hours of pain yesterday even if I would otherwise have had *no* pain yesterday. As Miller notes, at least one philosopher is on record predicting a similar intuition would be widely shared [Dorsey 2018: 1910], and empirical evidence has indicated that laypeople are still future-biased (though admittedly less so) when they sincerely believe they can affect the amount of pain in their pasts [Latham et al. 2021]. In fairness, Miller does argue that this evidence is inconclusive, but in the absence of additional studies it seems reasonable to assume for now

that the wider population would not share her intuitions, and accordingly that the bias toward the personal future is, indeed, what explains our reactions to the earlier cases.

Some readers may remain unconvinced by arguments which appeal to time-travel. Unfortunately, it is only in such cases that the personal and external futures diverge. There is, however, the following indirect argument for my conclusion. The bias toward the future is often thought to be intimately associated with the bias toward the near. And it would seem that an experience's proximity to the present in external time also does not matter to us.

Compare:

> *Short stasis*: You will enter a deep freeze, during which you will be completely unconscious for one day. It will seem to you as though no time has passed at all. Upon your revival, you will undergo ten hours of excruciating pain.
>
> *Long stasis*: You will enter a deep freeze, during which you will be completely unconscious for a hundred years. It will seem to you as though no time has passed at all. Upon your revival, you will undergo ten hours of excruciating pain.

As before, there may be contingent reasons for caring differently about these cases. However, restricting our focus to the pain itself, I predict almost everyone would be indifferent between them. If so, this provides some support to the general claim that it is personal time, not external time, which lies behind our prudential intuitions. It is therefore some circumstantial evidence that future bias concerns personal time rather than external time.

3. Asymmetries in the Grounds of Egoistic Concern

I will proceed on the assumption that we are biased toward the personal future. This suggests that to defend future bias, rather than looking to the temporal properties of past and future events, we should look instead to the relations we stand in to our past and future selves. (In what follows, by 'past' and 'future' I will mean 'personal past' and 'personal future'.) In this section, I want to explore how the proponent of future bias might defend her position as part of the general relation which grounds egoistic reasons over time. Specifically, I will propose that that relation is *asymmetric between person-stages*: typically, at a given time the degree to which a person stands in that relation to her future self is greater than the degree to which her future self stands in it to her.

Ordinarily, to determine the strength of my egoistic reason to care about some future experience, it is enough to know two things: how good or bad that experience will be for the person who has it, and whether that person will be me. Sometimes, however, these two considerations fail to capture our prudential intuitions, as the following two examples show. First, many of us believe that we begin to exist as foetuses, if not earlier in prenatal development. Yet we do not take the death of a foetus to be a profound tragedy—at least not for the foetus—even though the nearer death occurs to the beginning of one's existence, the more good life one loses by dying [McMahan 2002: 78]. Second, most people do not think that to gradually develop dementia would literally cause their non-existence. And yet those same people may well take their present egoistic reason to care what will happen to them in old age to be diminished if it will happen to them after they are severely demented—having, as they will, a very different personality and virtually no memories of their past life.

To account for our intuitions about these cases, it is useful to consider them in terms of what Jeff McMahan has called a *time-relative interest* [McMahan 2002: 65–74]. According to the Time-Relative Interest Account, the strength of one's present prudential reason to care about some event depends not only on (1) the value of the event, positive or negative, but also on (2) the degree to which the *egoistic concern relation* holds between oneself now and the individual whose welfare it affects. To capture our judgments about foetal death and future dementia, the Time-Relative Interest Account denies that *identity* is what grounds or constitutes the egoistic concern relation. Although the foetus would be one and the same individual as a future adult with a long and worthwhile life, and the pre-dementia person is the same individual as the one who eventually lives with severe dementia, these individuals are now at most weakly connected to their future selves in the way that grounds their prudential or egoistic interests in what will happen to them in the future.

I can now state my own proposal more fully. Previous discussants of future bias have focused primarily on (1)-type justifications. They have asked whether a pain's location in time could affect its *value*. But whether an event is in an individual's personal past or future is relative to a given time, so it is utterly unclear what intrinsic property of events could affect their value in the way that tracks what actually matters to us. I propose that the defender of future bias search instead for a (2)-type justification. To justify having greater concern for the personal future than the personal past, she should try to show that, at a given time, we typically stand in the egoistic concern relation to our future selves to a greater degree than we stand in it to our past selves.

As I say, there is good independent reason to accept that personal identity is not the basis of the egoistic concern relation. But it must be said that if identity *were* the basis, then the strategy for justifying future bias I am proposing would not be very promising. That is because identity is symmetric. As such, taking personal identity to ground egoistic concern over time means that whatever relation underwrites one's present egoistic reason to care about *future* person-stages also underwrites one's present egoistic reason to care about *past* person-stages. That is true, moreover, even if we are reductionists about personal identity over time, believing that it ultimately consists in more particular relations which can be stated in an entirely impersonal way. For even if those more particular relations are not themselves symmetric—as those involved in anticipation, intention, and memory are not—since identity is symmetric, the converse of any relation that is necessary or sufficient for it must also be necessary or sufficient for it. So, while the converse of 'remembers having the experiences of' is the entirely different relation 'is having experiences that will be remembered by', if identity forms the basis of egoistic reasons over time, then if one of those relations is necessary or sufficient for having reasons of egoistic concern, then the other is, too.

Since personal identity is symmetric, it is easy to take for granted that the egoistic concern relation must be, too. But of course that does not follow. For the purpose of illustration, assume that at t_1 (which is now) I should be egoistically concerned with what happens to me at t_2 when and because at t_2 I will remember the experiences that I am now having. It remains an open question whether at t_2 I should be equally concerned with what happened to me at t_1. In this way, those who reject an identity account of the egoistic concern relation are free to replace it with a relation that is asymmetric between person-stages. If we could establish that the relation is indeed asymmetric, and in particular asymmetric in the direction that warrants giving greater consideration to future person-stages than to past ones, then we would be justified in being biased toward the personal future.

Such an attempt would be rendered most plausible if we could point to some particular component of the egoistic concern relation which is itself asymmetric. Ideally, we would also have a freestanding argument that the relation is asymmetric in the requisite direction—an argument which makes no appeal at all to the fact that we are biased toward the personal future. Presently I shall attempt to do both, but first let me briefly say why there may already at least some reason to accept this way of justifying future bias.

Ultimately, to discern our considered beliefs about the grounds of egoistic concern over time, we have to reflect on cases. And per-

haps we should think of cases like *Your Past or Future Operation* and *To the Middle Ages* as members of that diverse family of examples whose purpose is to lay bear our prudential intuitions. That family includes radical personality changes, cloning, brain damage, brain transplants and divisions, and so on. If our intuitions about those cases support the move to some particular account of the egoistic concern relation—one which gives pride of place to psychological connections, for example—then perhaps the fact that we are biased toward the personal future supports the move to an asymmetric egoistic concern relation.

With that aside, I will turn now to the question of what component of the egoistic concern relation is a good candidate for being asymmetric between the personal past and future. My tentative suggestion is that the asymmetry is to be found in the causal or counterfactual dependence that is necessary for the relation to obtain between person-stages in the first place.

To see what I am getting at, consider the following case:

> *Vaporisation-Replacement*: You somehow come to learn that tonight, not only will you be vaporised in your sleep, but a random duplicate of you will later pop into existence, replacing you in your bed. This duplicate will not be created by, for example, scanning your body and reproducing it from new matter—there is no counterfactual dependence whatsoever. It is just a quirk of chance that will happen to occur.

To many, this prospect would seem like a deep misfortune. While perhaps not quite as bad as ordinary death, it seems far closer to it than to ordinary survival. To account for this intuition, nearly all candidates for the egoistic concern relation have incorporated some notion of continuity over time, understood as counterfactual or causal dependence (or chains of such dependence) connecting one person-stage to another. For some writers, this continuity consists in the dependence of mental states at one time on mental states at another, the way a memory depends on an experience [Parfit 1971, 1984: 261–302; Perry 1976; Shoemaker 1984]. For others, it consists, or also consists, in the physical and functional continuity of certain parts of the brain [Unger 1990; McMahan 2002]. But we can ignore these differences here.

Now consider a variation of *Vaporisation-Replacement*. This time you learn that you are the replacement. Some duplicate of you was vaporised last night and randomly replaced by you in the bed you woke up in this morning. This news would no doubt be unsettling, but if you are like me, it is simply not comparable to the original version. Because you are biased toward the future, you care far less

about the past life you have lost out on than you cared about the future life you are to lose out on in the initial version of the case.

To account for these two contrasting reactions, a defender of future bias can amend the continuity requirement for egoistic concern over time. On the received picture of continuity, which I call the *Two-Directional View*, the chains of dependence that are necessary for egoistic reasons over time can run in either the 'downstream' or 'upstream' direction. But giving up identity as the basis for egoistic concern means we are free to instead adopt the *One-Directional View*. This position holds that the chains of dependence run in the downstream direction only. On both views about continuity, *future* vaporisation severs the egoistic concern relation between you and your future self, so the good life that will be lived by your replacement is no longer worth wanting. But according to the One-Directional View only, the relation connecting you to your *past* self never obtained in the first place. Thus, when it turns out that you are a replacement, you did not lose something you had reason to want in the first place.

Importantly, the One-Directional View would seem to recommend *absolute* future bias. Someone who is absolutely future-biased assigns not merely *less* egoistic significance to her past experiences, but no significance at all [Greene and Sullivan 2015: 961–2]. While some may consider this a virtue of the One-Directional View, empirical research indicates that most people among the general population do not discount past pleasures absolutely [Greene et al. 2021b].

Those troubled by absolute future bias may be partly assuaged by emphasising that the One-Directional View is strictly intended as a view about egoistic or prudential reasons. It does not rule out having reasons of, say, sympathy for one's past self—any more than a mainstream psychological continuity account rules out having reasons of sympathy for close friends with whom one is not psychologically continuous.

Alternatively, the defender of future bias could avoid absolute future bias by eschewing the One-Directional View but upholding the broader claim that the egoistic concern relation is asymmetric in the way that warrants greater concern for one's personal future. Whatever that relation is, it will plausibly come in degrees. So we could simply say that at a given time, someone typically bears it to past person-stages to merely a lesser degree than to future ones.

Still, justifying the bias toward the future on the basis of an asymmetric egoistic concern relation, without so much as pointing to a component of that relation that is itself supposedly asymmetric, may be too nearly circular to sway someone not already predisposed to endorse future bias. So I want to explore a freestanding argument that the egoistic concern relation is asymmetric in the direction that warrants having more egoistic concern for future person-stages than

past ones—an argument which makes no reference whatsoever to the fact that we *are* future-biased.

The argument proceeds from yet another case which has convinced many people that identity is not what grounds egoistic reasons over time. The case is due to David Wiggins [1967: 53] but developed most famously by Parfit [1971, 1984: 253–66]. In it, we imagine that the hemispheres of one person's brain are divided and separately transplanted into the brainless bodies of two others. For good measure, we stipulate that each of the resulting people believes he is the original pre-division person, remembers living his life, has the same personality, and is psychologically continuous with him in every other respect [Parfit 1984: 254]. Call the two post-division individuals 'Lefty' and 'Righty' and the original individual 'Wholly'.[5] This case is thought to show that identity is not necessary for reasons of egoistic concern over time because it has seemed to many that Wholly has egoistic reason to care what happens to each of Lefty and Righty. However, given that identity is transitive, and Lefty and Righty do not seem to be the same person, it seems that Wholly is identical to neither.

The first premise of my independent argument for the asymmetry of the egoistic concern relation is that Wholly can stand fully in that relation to each of Lefty and Righty, at least as they are shortly after the division. This follows from two widely accepted sub-premises. First, if just one of Lefty and Righty had existed—because the other hemisphere was destroyed due to a stroke, say—then Wholly could have stood fully to that individual in the way that grounds Wholly's egoistic reasons to care what happens to him, at least that individual is shortly after the division. The other sub-premise is that the egoistic concern relation is intrinsic, such that the extent to which one person-stage stands in it to another is not diminished by his also standing in it to someone else [Parfit 1984: 263].

The second premise is this. Provided one person (or person-stage) *X* stands fully to another *Y* in whatever way grounds *X*'s present egoistic reasons to care about *Y*, then *X* should give about as much weight to *Y*'s egoistic reasons as *Y* should. Speaking a bit loosely, this principle would tell us that provided *X* should be fully egoistically concerned for *Y*, then *X* should also care about *Y*'s future pain about as much as *Y* should.[6]

---

[5] The name 'Wholly' appears first in Gustafsson [2019: 2311].

[6] I emphasise that this principle is intended to apply only when *X* stands *fully* in the egoistic concern relation to *Y*, such that *X* rationally should give as much weight to *Y*'s pain as he should give to his own. Suppose, by contrast, that *X* stood in that relation only partially to *Y*. If *Y* in turn stood in it only partially to *Z*, *X*'s egoistic reason to care about *Z*'s pain plausibly would be much weaker than *Y*'s own reason to care about it.

Supposing that we grant these two premises, the rest of the argument is straightforward. By the first premise, we stipulate that Wholly stands fully to Lefty in the way that that grounds Wholly's egoistic reason to care what happens to Lefty. Now assume for contradiction that the egoistic concern relation is symmetric. Then Lefty should likewise have full egoistic concern for Wholly. From our second premise, it follows that Lefty's egoistic reason care about someone's pain is about as strong as Wholly's egoistic reason to care about it. Since by assumption Wholly is fully related to Righty, it then follows that Lefty should give about as much weight to Righty's pain as Lefty should give to his own impending pain. But that result is absurd. Imagine being Lefty and learning that either you or Righty will soon undergo some agony. It is rather difficult to believe, I predict, that from a purely egoistic perspective, your interest in averting Righty's pain is anywhere near as strong as your interest in averting your own.

To avoid that conclusion, we must reject one of our assumptions. The first premise holds that Wholly can stand fully in the egoistic concern relation to each of his two offshoots. While this premise is not beyond dispute, it is, I think, highly intuitive. Imagine being Wholly. You initially believe that just one hemisphere of your brain will survive and be transplanted, and that the resulting person will immediately thereafter undergo some agonising pain. Does that pain really seem less bad when you are now informed that the other hemisphere of your brain will actually survive?

The second premise says that when we are fully related to another person (or person-stage) in the way that grounds our present egoistic reason to care what happens to them, then we should give about as much weight to their self-regarding reasons—their egoistic interests—as they should. While that principle is not incontrovertible either, it must be admitted that its falsity would be surprising. Having a given degree of egoistic concern about what happens at t to someone does not *imply* having the same degree of concern for the egoistic interests they then have in what happens at times other than t, but it is hard to see the two attitudes rationally diverging substantially. Suppose I could somehow stand fully to you in the way that grounds my egoistic reason to care what happens to you now. I then have no less egoistic reason to care what happens to you as I have to care what happens to me. If you should be greatly concerned about your future pain, it would be strange to think that I should not have a great deal of concern about it, too.

We would do better, I think, to deny the symmetry of the egoistic concern relation. More exactly, we should deny any account of that relation which is not asymmetric in the direction that favours the personal future, since an asymmetry in the opposite direction would

also yield the result that Lefty should be about as concerned with Righty's pain as with his own.

4. Our Attitudes Toward Other People's Past and Future

Justifying future bias on the basis of the egoistic concern relation has another advantage. It resolves the tension in what Parfit described as 'a surprising asymmetry in our concern about our own, and other people's pasts' [1984: 182]. To see the asymmetry Parfit was getting at, first consider:

> *My Mother's Future Pain*: I am an exile from some country, where I have left my widowed mother. Though I am deeply concerned about her, I very seldom get news. I have known for some time that she is fatally ill, and cannot live long. I am now told something new. My mother's illness has become very painful, in a way that drugs cannot relieve. For the next few months, before she dies, she faces a terrible ordeal [1984: 181].

Here, I would be saddened that my mother will die soon. But I already knew that. The news that she must also suffer greatly before dying would deeply distress me on her behalf.

Now consider a further development:

> *My Mother's Past Pain*: A day later, I am told that I have been partly misinformed. The facts were right, but not the timing. My mother did have many months of suffering, but she is now dead [1984: 181].

Parfit put it to his reader whether this new knowledge would make her relieved for her mother's sake. I share his reaction [1984: 181–2] that it would not. The revelation that my mother's agony is now in the past does virtually nothing to mitigate my distress about her ordeal. More generally, like others [Hare 2008, 2013; Brink 2011: 378–79; Dougherty 2015: 3; Sullivan 2018: 123–6], I find myself disposed to be *time-neutral* about other people's pain.[7]

---

[7] Hare [2008, 2013], however, reports that although he would be time-neutral on another person's behalf when she is far away, he would be future-biased on her behalf when she is nearby and her pain salient. But it may be our intuitions about the far-away case which are relevant to the question of whether we actually endorse other-regarding time-neutrality *per se*. When someone is up close, we may be inclined to base what we want to happen to her on what we believe that she herself would want—an ethical consideration not present in the first-personal case. I expand on this point in footnote 8.

This appears to threaten the justifiability of future bias in the first-personal case, on the grounds that whatever considerations could warrant caring less about one's own past pain would seem to warrant caring less about the past pain of others. I call this the challenge from *other-regarding time-neutrality*.

Now, the defender of (first-personal) future bias could evade the challenge by simply denying that we should be time-neutral, rather than future-biased, when it comes to the pain of others. And this may be a more attractive position than Parfit thought. For although philosophers' intuitions seem to favour other-regarding time-neutrality, in several studies, most participants who were asked to consider the experiences had by a third party wanted his good ones to be in his future and his bad ones in his past [Greene et al. 2021a, 2021b, forthcoming].

It is debatable whether that evidence ultimately helps the defender of future bias avoid the challenge.[8] But in fact the opponent of future bias might not even need to appeal directly to our intuitions about a case like Parfit's. For there exists the following powerful indirect case for other-regarding time-neutrality: it is a short step from the claim that we should be future-biased on behalf of those whose welfare we are concerned to promote, to the conclusion that we should apply future-directed distributive principles when multiple other people's interests conflict. Consider:

> *Two Patients and a Drug*: Amy and Betty must each undergo two intensely painful operations spaced out over several days. Amy has already had her first operation, which took place yesterday and involved 5 hours of pain. Betty's first operation will take place later today, and will involve 4 hours of equally intense pain. Both Amy and Betty will have their second operations tomorrow, each of which, if nothing is done, will involve 5 hours of pain. We are now made aware of a single dose of a drug, which, provided it is administered immediately, will

[8] As Greene et al. [2021a: 159–60] themselves note, there is an apparent contradiction between their findings and those of Caruso, Gilbert, and Wilson [2008], most of whose participants *did* display other-regarding time-neutrality. Greene et al. offer the following hypothesis. While participants in the Caruso et al. study were prompted to consider the fate of an unidentified third party, the Greene et al. vignettes contained sufficiently rich biographical information that their participants identified with the third party and thereby adopted his (presumably future-biased) preferences. However, what another person prefers might well constitute an additional ethical consideration which partially determines what we want to happen to them—a consideration that is absent in the first-personal case (see Parfit [1984: 182]; Greene et al. [2021a: 159–60]). Since the challenge from other-regarding time-neutrality only requires that we display such time-neutrality when other things are equal, the Caruso et al. results may well be sufficient to press that challenge to first-personal future bias.

allow us to safely reduce the duration of the second operation by 1 hour for one patient.

I suspect that almost everyone would think we should give the drug to Amy, who faces more total pain.[9] Yet if someone else's pain matters more when and because it is in her future, then many widely-endorsed principles of distributive ethics—for example equality; priority to benefitting the worse off—would seem to recommend giving the drug to Betty, who is yet to have her first operation. After all, she faces nine hours of intense future pain, whilst Amy faces only five.

In the end, I do not wish to take a rigid stance on whether we should be future-biased when other people's interests are at stake. But I do think the proponent of future bias should be prepared to accept that her position does not extend to the other-regarding case. Now, for those who would defend future bias on the basis of deep metaphysical or physical differences between the past and the future, this *self-other* asymmetry is very difficult to uphold. The properties of external time touch all lives alike. So if *my* past pain should matter less to me now because of its location in external time, it is hard to see why *your* past pain should not also matter less to me by the same token.

If, on the other hand, we justify future bias on the basis of an asymmetry within the egoistic concern relation, any tension between our first-personal and other-regarding intuitions dissolves. The facts that ground my egoistic reason to care more about pain I will have tomorrow than pain I had yesterday do not similarly ground a reason to care more about *your* pain tomorrow than your pain yesterday. The present stage of me is no more connected to the present stage of you than it is to any other stage of you; put otherwise, it is equally unconnected to them all. Hence, I would have no special reason to care about how you are faring now by comparison with how you fared yesterday, or will fare tomorrow. A justification that appeals to an asymmetry in the egoistic concern relation therefore avoids another of the major challenges that have been raised against future bias.

5. Summary

Parfit once wrote that 'if time's passage is an illusion, temporal neutrality cannot be irrational', by which he meant that we could not

[9] Dennis McKerlie [1989: 480] also rejects future-directed egalitarianism on the grounds that it can generate conflicting imperatives from one time to the next, but justice should not instruct us to undo what it earlier demanded.

have more reason to prefer, for its own sake, that an event is in the future rather than the past [Parfit 1984: 184]. This, I have said, is a mistake. Our bias toward the future corresponds not to the locations of good and bad things in external time, but with their locations relative to the present within our own lives.

Those seeking to justify future bias would therefore do well to look instead to the relation we now stand in to past and future stages of ourselves. In particular, they might look to the relation that underwrites our present egoistic reasons to care about what happens to those other stages. Philosophers have overlooked the possibility that at a given time, a person typically stands in that relation to her past self to a lesser degree than she does to her future self. That claim, which has some independent warrant, would justify the kind of future bias human beings actually exhibit.

Accounting for future bias in this way would also prise apart our reasons to care about our own pasts and futures from the reasons we have to care about those of others. There would be no tension, therefore, in having time-neutral preferences about other people's circumstances, but future-biased preferences about our own.[10]

References

Brink, David 2011. Prospects for Temporal Neutrality, in *The Oxford Handbook of Philosophy of Time*, ed. C. Callender, Oxford: Oxford University Press: 353–81.

Caruso, Eugene, Daniel Gilbert, and Timothy Wilson 2008. A Wrinkle in Time: Asymmetric Valuation of Past and Future Events, *Psychological Science* 19/8: 796–801.

Cockburn, David 1998. Tense and Emotion, in *Questions of Time and Tense*, ed. R. Le Poidevin, Oxford: Oxford University Press: 77–91.

Dorsey, Dale 2017. Future-Bias: A (Qualified) Defense, *Pacific Philosophical Quarterly* 98/1: 351–73.

[10] For their valuable input, I am grateful to Marius Backmann, Tomi Francis, Johan Gustafsson, Kacper Kowalczyk, Michal Masny, Jeff McMahan, Véronique Munoz-Dardé, Michael Otsuka, Aidan Penn, Nicolas Rimell, and Bastian Steuwer, as well as two anonymous referees and an Associate Editor for this journal.

Dorsey, Dale 2018. Prudence and Past Selves, *Philosophical Studies* 175/8: 1901–25.

Dougherty, Tom 2011. On Whether to Prefer Pain to Pass, *Ethics* 121/3: 521–37.

Dougherty, Tom 2015. Future-Bias and Practical Reason, *Philosophers' Imprint* 15/30: 1–16.

Garrett, Brian 1988. 'Thank Goodness That's Over' Revisited, *Philosophical Quarterly* 38/151: 201–5.

Greene, Preston, Andrew Latham, Kristie Miller, and James Norton 2021a. Hedonic and Non-Hedonic Bias Toward the Future, *Australasian Journal of Philosophy* 99/1: 148–63.

Greene, Preston, Andrew Latham, Kristie Miller, and James Norton 2021b. On Preferring that Overall, Things are Worse: Future-Bias and Unequal Payoffs, *Philosophy and Phenomenological Research*, URL = <https://doi.org/10.1111/phpr.12819/>

Greene, Preston, Andrew Latham, Kristie Miller, and James Norton. How Much Do We Discount Past Pleasures? *American Philosophical Quarterly,* forthcoming.

Greene, Preston and Meghan Sullivan 2015. Against Time Bias, *Ethics* 125/4: 947–70.

Gustafsson, Johan 2019. Non-Branching Personal Persistence, *Philosophical Studies* 176/9: 2307–29.

Hare, Caspar 2007. Self-bias, Time-Bias, and the Metaphysics of Self and Time, *Journal of Philosophy* 104/7: 350–73.

Hare, Caspar 2008. A Puzzle About Other-Directed Time-Bias, *Australasian Journal of Philosophy* 86/2: 269–77.

Hare, Caspar 2013. Time – The Emotional Asymmetry, in *A Companion to the Philosophy of Time*, eds. A. Bardon and H. Dyke, Malden, MA: Wiley-Blackwell: 507–20.

Heathwood, Chris 2008. Fitting Attitudes and Welfare, in *Oxford Studies in Metaethics, Volume 3*, ed. R. Shafer-Landau, Oxford: Oxford University Press: 47–73.

Kamm, F. M. 1993. *Morality, Mortality, Volume I: Death and Whom to Save from It*, New York, NY: Oxford University Press.

Karhu, Todd 2019. *Matters of Life and Death*, Doctoral thesis, London School of Economics and Political Science, URL = <http://etheses.lse.ac.uk/3917/>

Latham, Andrew, Kristie Miller, James Norton, and Christian Tarsney 2021. Future Bias in Action: Does the Past Matter More When You Can Affect It? *Synthese* 198/12: 11327–49.

Lee, Ruth, Christoph Hoerl, Patrick Burns, Alison Fernandes, Patrick O'Connor, and Teresa McCormack 2020. Pain in the Past and Pleasure in the Future: The Development of Past–Future Preferences for Hedonic Goods, *Cognitive Science* 44/9: 1–44.

Lewis, David 1976. The Paradoxes of Time Travel, *American Philosophical Quarterly* 13/2: 145–52.
McKerlie, Dennis 1989. Equality and Time, *Ethics* 99/3: 475–91.
McMahan, Jeff 2002. *The Ethics of Killing: Problems at the Margins of Life*, New York, NY: Oxford University Press.
Miller, Kristie 2021. What Time-travel Teaches Us About Future-Bias, *Philosophies* 6/2: 1–18.
Parfit, Derek 1971. Personal Identity, *The Philosophical Review* 80/1: 3–27.
Parfit, Derek 1984. *Reasons and Persons*, Oxford: Clarendon Press.
Perry, John 1976. The Importance of Being Identical, in *The Identities of Persons*, ed. A. Rorty, Berkeley: University of California Press: 67–90.
Prior, Arthur 1959. Thank Goodness That's Over, *Philosophy* 34/128: 12–17.
Scheffler, Samuel 2021. Temporal Neutrality and the Bias Toward the Future, in *Principles and Persons: The Legacy of Derek Parfit*, eds. J. McMahan, T. Campbell, J. Goodrich, and K. Ramakrishnan, Oxford: Oxford University Press: 85–114.
Schlesinger, George 1976. The Stillness of Time and Philosophical Equanimity, *Philosophical Studies* 30/3: 145–59.
Shoemaker, Sydney 1984. Personal Identity: A Materialist's Account, in *Personal Identity: Great Debates in Philosophy*, eds. S. Shoemaker and R. Swinburne, Oxford: Blackwell: 67–132.
Sullivan, Meghan 2018. *Time Biases*, Oxford: Oxford University Press.
Tarsney, Christian 2017. Thank Goodness That's Newcomb: The Practical Relevance of the Temporal Value Asymmetry, *Analysis* 77/4: 750–59.
Unger, Peter 1990. *Identity, Consciousness, and Value*, New York, NY: Oxford University Press.
Wiggins, David 1967. *Identity and Spatio-Temporal Continuity*, Oxford: Blackwell.
Yehezkel, Gal 2014. Theories of Time and the Asymmetry in Human Attitudes, *Ratio* 27/1: 68–83.