

Time-biased Attitudes and Time-biased Preferences

Brad Skow

Version: April 24, 2008

1 What is the Bias Toward the Future?

In part two of *Reasons and Persons*, “Rationality and Time,” Derek Parfit introduced us to the bias toward the future. What is the bias toward the future?

There are actually two biases that Parfit calls “the bias toward the future.” We might have biased attitudes, or biased preferences. Parfit does not clearly distinguish between them.¹ My main goal in this paper is to sort out what these different biases amount to, and what the neutral state these biases depart from looks like. I conclude by saying something about the relationship between the biases. I claim that our having biased preferences explains our having biased attitudes.

2 Biased Attitudes

Parfit first introduces the bias toward the future in this passage:

...the bias towards the future applies most clearly to events that are in themselves pleasant or painful. The thought of such events affects us more when they are in

¹Most writers on this topic focus only on one or the other of these biases. Moller (2002), for example, defends the rationality of biased attitudes, but says nothing about biased preferences. Kagan (1986), on the other hand, is clearly talking about biased preferences, but says nothing about biased attitudes.

the future rather than the past. Looking forward to a pleasure is, in general, more pleasant than looking back on it. (Parfit 1984, page 160)

Someone who finds looking forward to a pleasure more pleasant than looking back on it has *biased attitudes*. Before saying more precisely what it is to have biased attitudes, we need to know what kind of pleasures and pains Parfit is talking about. Perhaps he is talking about sensory pleasures (like the taste of chocolate) and sensory pains (like the experience you would have if someone stuck a needle under your fingernail). Perhaps, on the other hand, he is talking about *attitudinal* pleasures, and attitudinal pains. Attitudinal pleasure is a propositional attitude. We ascribe this attitude to someone when we say things like, “Joe takes pleasure in the fact that Fermat’s last theorem has been proven.”² This gives us four possible readings of the claim that looking forward to a pleasure is pleasant. The most plausible reading is

- (1) When I think about a future sensory pleasure of mine, I take pleasure in the fact that I will experience it.

Now I’m in a position to characterize biased attitudes. As a first pass, say that someone *S* has perfectly biased (toward the future) attitudes iff he satisfies these four conditions:

- (i) When *S* thinks about a future sensory pleasure of his, he takes pleasure in the fact that he will experience that pleasure;
- (ii) When *S* thinks about a future sensory pain of his, he takes pain in the fact that he will experience that pain;
- (iii) When *S* thinks about a past sensory pleasure of his, he takes no pleasure or pain in the fact that he experienced it;
- (iv) When *S* thinks about a past sensory pain of his, he takes no pleasure or pain in the fact that he experienced it.

²Fred Feldman (1997) argues persuasively that attitudinal pleasure is not to be analyzed in terms of sensory pleasure.

But this isn't quite right, for three reasons.

First: certainly, someone who satisfies (i)-(iv) does not have the same attitudes toward pleasures and pains he has yet to experience that he has toward pleasures and pains he has already experienced. But nothing in (i)-(iv) ensures that the difference in when these sensations occur is the *source* of the asymmetry in *S*'s attitudes. Maybe I'm having some difficult dental work done. I just had some surgery yesterday, and tomorrow I will have some more, by a different doctor. The first doctor was a nice guy, a friend of mine, while the doctor I will see tomorrow is a jerk. I might take pain in the fact that I will experience pain tomorrow, because I take pain in the fact that the money I will pay for the surgery will go to the doctor I dislike. But I don't take this attitude toward the pain I experienced yesterday, because the doctor I saw then is a friend. Here I have asymmetric attitudes, but the fact that I have already experienced one of the painful episodes, while I have yet to experience the other, is not the source of the asymmetry.

These instances of propositional pain are extrinsic, or derived: I take pain in the fact that I will experience some sensory pain because, or in virtue of the fact that, I take pain in the fact that my money will go to a jerk. If we restrict (i)-(iv) to episodes of intrinsic, or underived, propositional pleasure and pain, then we avoid this kind of interference.

Second: according to (i)-(iv) the bias toward the future is an asymmetry in our attitudes toward future and past *sensory pleasures* and *sensory pains*. But why this limitation to sensory pleasures and pains? What if someone isn't particularly interested in sensory pleasure or pain? He likes to meditate in a quiet, darkened room, having as few sensory experiences as possible. That is when he is most content. This guy's attitudes might be biased toward the future: he might take great pleasure in the fact that he'll have some time to meditate tomorrow, and be indifferent to the fact that he got to meditate yesterday. But he does not satisfy (i)-(iv).

That's right: although Parfit seems to limit the bias toward the future to asymmetric attitudes towards future and past sensory pleasures and pains, a better definition of "biased attitudes" wouldn't have this limitation. Still, I'm going to stick with the focus on sensory pleasures

and pains. The difference between the restricted version of the bias and the more general version doesn't matter to what I want to say.

Third: if someone fails to satisfy (i)-(iv) it does not follow that his attitudes are completely unbiased. Someone might, for example, take some pleasure in the fact that he has already experienced some episodes of pleasure, but take *more* pleasure in the fact that he is going to experience some other episodes of pleasure; or he might satisfy (i) and (iii) but have symmetric attitudes toward future and past pain; or he might take *pleasure* in the fact that he's made it through some episode of pain (thinking to himself, "Thank goodness that's over!"); or he might on some occasions take more pleasure in pleasures he has yet to experience than in those he already has, and on other occasions not. Someone can exhibit this bias to a greater or lesser degree. In addition, being biased toward the future is only one of many ways of having time-biased attitudes. There is also, for example, the bias toward the present. Someone with this bias would only enjoy pleasant experiences while they are present, and would take neither pleasure nor pain in them before or after they occurred.

So what would it be to have attitudes that are not temporally biased in any way? The guiding idea is that an unbiased person's attitudes toward the pleasures and pains he experiences are the same, whether those pleasures and pains are in his future or his past. Perhaps this is a way to make this idea more precise. Suppose that E is some sensory pleasure I know that I will experience, and E^* is some sensory pleasure in my past that I know I already had. Suppose also that E and E^* are exactly alike in intensity and duration. We might define "unbiased attitudes" so that if I am unbiased, I always take the same amount of pleasure in the fact that I will experience E as I do in the fact that I have already experienced E^* . That's what we want our definition to say about this particular case. The definition is just a generalization:

(UA) Someone has temporally unbiased attitudes iff, for any sensory pleasure of his E (that he knows he experiences) there is some amount of pleasure such that at each time he takes that amount of pleasure in the fact that he will experience (or has experienced, or is

experiencing) *E*, and similarly for pain.³

3 Biased Preferences

When Parfit first introduces the bias toward the future, he's talking about biased attitudes. A few pages later Parfit mentions the bias toward the future again; but there he appears to be talking about a different phenomenon. In a famous passage, he describes a scenario in which he must undergo some painful surgery without anesthetic, after which the doctors will put him to sleep and give him an amnesiac drug. He wakes up and does not know whether he has had the surgery yet. The nurse doesn't know either—she is not sure which patient he is. She tells him that either he has already undergone the operation, or he still has yet to undergo it. If he is the patient who already had the surgery, then it took ten hours. If he is the patient who has not yet had the surgery, then no complications are expected and it will take about one hour. Parfit reports: “it is clear to me which I prefer to be true. If I learn that the first is true, I shall be greatly relieved” (page 166).

Here “the bias toward the future” refers to a feature of someone's preferences (or wants, or desires), rather than to a feature of his episodes of attitudinal pleasure and pain.⁴ Parfit prefers being the first patient to being the second patient. The life he prefers is worse overall (it contains more pain), but has a better future (the extra pain is in the past). So in this case Parfit's preferences are tracking future welfare (how well someone is doing from now forward), rather than overall welfare (how well someone is doing throughout his entire life). Someone who exhibits this kind of bias toward the future has *biased preferences*.

³Note that note that the restriction I made earlier to episodes of underived attitudinal pleasure is still in force. Timeless, Parfit's example of someone with unbiased attitudes, fits this characterization (page 174).

⁴I think that having biased attitudes is one thing, and having biased preferences is another. Not everyone will agree with this. If you think that attitudinal pleasure is to be analyzed in terms of desire or preference, then (depending on the details of your theory) you might end up thinking that the biases are the same. (Heathwood (2005) defends such an analysis.) But I am going to assume this is wrong.

The concept of biased preferences, then, depends on the concept of individual welfare, of how good someone's life is for the person who lives that life. It also presupposes that we can make sense, not only of the welfare level of someone's entire life, but also of the welfare level of parts of someone's life—in particular, the future part of his life, the part that runs from now until his death.

One more quick point of clarification before proceeding: this first statement of what it is to have biased preferences isn't quite right. Parfit's preferences aren't really tracking future welfare; they're tracking Parfit's *beliefs* about future welfare. Parfit might have false or incomplete beliefs about what happens to the patients. Perhaps it is actually the first patient who has the worse future, due to catastrophes that will later befall him that Parfit doesn't know about. Parfit still prefers being the first patient, since he believes the first patient to have the better future. Or Parfit might have false beliefs about what makes a life go well. Someone in Parfit's position might think that pain actually makes a life go better, and so prefer being the second patient to being the first. (If it is hard to imagine someone who is not a masochist preferring being the second patient, that is because it is hard to imagine someone who thinks that pain makes a life go better.) From now on, I will leave these qualifications tacit.

Parfit's preferences are *de se*. They are not just preferences about which possible world is actual, but also preferences about which person in the actual world is *him*. We can imagine the nurse has given Parfit complete information about the two patients: she tells him that the first patient is Derek Parfit, and that Parfit has already had his operation; that the second patient is John Doe, and that Doe has not yet had *his* operation; and so on. So Parfit knows, in relevant respects, which world is the actual world. He knows, in particular, that it is Parfit, and not Doe, who has already had his operation. But since Parfit does not know that he is Derek Parfit, there is still something he doesn't know. He doesn't know which of those two people is him. And that is what his preference is about: he is not preferring that one possible world rather than another be actual; he is preferring that he have one location within the actual world, rather than another.

A simple modification of the case shows that one may manifest the bias toward the future

by having *de se* preferences about which time is *now*. Suppose, as before, that Parfit wakes up and does not know whether he has had the surgery yet. The nurse doesn't know either—but this time, she knows that he is Derek Parfit. She doesn't know whether he has had the surgery yet because she has lost track of the time. She remembers that the surgery was scheduled from ten to eleven o'clock, so she tells him that if it is before ten, he has not yet had the surgery, and if it is after eleven, he has. Parfit knows (in relevant respects) which world is actual. Still, he prefers that it now be after eleven, because the part of his life from eleven onward has higher welfare than the part from ten onward. For another example, suppose that I have just had a really good year, and I prefer it to now be one year ago. Then I am biased toward the future: my life is equally good, with respect to welfare, as itself; but the part of my life from a year ago on has a higher level of welfare than the part of my life from now on.

Parfit's preference in this case is a particular instance of biased preferences. But what is it to have biased preferences, generally? Someone with biased preferences is, well, biased. He might be more or less biased. His preferences might deviate from an unbiased state to a greater or lesser degree. What is this unbiased state? It is a state in which his preferences are insensitive to the "time coordinate":

(UP) If S prefers being person p_1 at t_1 in world w_1 to being person p_2 at t_2 in world w_2 , then for any times t_3 and t_4 , S prefers being p_1 at t_3 in w_1 to being p_2 at t_4 in w_2 .⁵

Someone who always prefers to be the person with the better life, and does not care whether that person has the better future, satisfies (UP). (And Parfit fails to satisfy (UP) exactly because he does not have this preference.) But that is not the only way to be unbiased; even people who always prefer to be the person with the worse life lack the bias toward the future.

Now I can state what it is to have perfectly biased preferences: S 's preferences are perfectly biased toward the future just in case:

(BP) S prefers being person p_1 in w_1 at t_1 to being p_2 in w_2 at t_2 just in case the part of p_1 's life

⁵This definition was proposed by Richard Chappell.

from t_1 until his death is better—has a higher level of well-being—than the part of p_2 's life from t_2 until his death.

As I said earlier, some people might have imperfectly biased preferences. Almost everyone shares Parfit's preference in the hospital case. But with respect to other cases they fail to satisfy (BP). I sometimes say that I wish I were sixteen again. People often tell me that *they* do not want that at all—"high school sucked," they tell me. They say this even though the part of their life from their sixteenth birthday forward has a higher level of welfare than the part of their life from now forward. If they had perfectly biased preferences, they would prefer to be sixteen again.

I've defined "perfectly biased preferences" with (some version of) hedonism about welfare in the back of my mind. Things get a little strange if some other theory of welfare is true. Some philosophers think there can be unexperienced evils. They think that if my friends ridicule me behind my back, then that is bad for me, even if I never know about it. If there can be unexperienced evils, then someone with perfectly biased preferences will prefer that they, too, be in the past. I'm confident that most of us prefer that our pains be in the past. But it's less clear to me that, supposing that there can be unexperienced evils, most of us prefer to have them in the past. Maybe the thing to say is that, if no mental state theory of welfare is true, we are not perfectly biased toward the future. On the other hand, maybe it means that I have mischaracterized biased preferences. Maybe I should give a definition that says, roughly, that I have biased preferences iff my preferences track future net balance of pleasure over pain, rather than future welfare. Parfit's hospital example is just as much an instance of this bias as it is of (BP). But I don't think there's much sense in wondering which generalization captures what is *really* going on in that case.

Being biased toward the future is not the only way to be temporally biased. I could be biased toward the present, preferring $\langle w_1, p_1, t_1 \rangle$ to $\langle w_2, p_2, t_2 \rangle$ just in case p_1 is doing better at t_1 than p_2 is doing at t_2 . Or I could be biased toward the near, preferring $\langle w_1, p_1, t_1 \rangle$ to $\langle w_2, p_2, t_2 \rangle$

just in case p_1 's "discount-adjusted future welfare level" is higher than p_2 's.⁶ (To compute someone's discounted-adjusted future welfare level at a time, multiply his welfare level at each future time by his "discount-rate"—a function that decreases as it goes further into the future. Then take the sum.) Or I might have future Tuesday indifference, preferring $\langle w_1, p_1, t_1 \rangle$ to $\langle w_2, p_2, t_2 \rangle$ just in case the Tuesdayless part of p_1 's future is better than the Tuesdayless part of p_2 's future.

Other biases don't have anything to do with time. Most of us are also biased toward the actual. I could have led a very different life, one in which I knew none of the people I actually know, including the people who are most important to me. That other life might also have been a better life than the life I'm leading. Still, I prefer my actual life to that other life.⁷

4 Comparing the Biases

I've said what the two biases are, and what it would be to completely lack either bias. I want to conclude by saying something about the connection between the two biases. Not only do we have both biases; I think that the fact that we have one of them explains why we have the other. Biased preferences explain biased attitudes.

In general, there is some connection between desire and pleasure. It is a commonplace that very often, when we get what we want, we take pleasure in that fact. (That is, if I desire that p , then when I learn that p , I take pleasure in the fact that p .)⁸ Now suppose I have biased

⁶Parfit's discussion of the bias toward the near occurs alongside his discussion of the bias toward the future. But his definition of the bias toward the near differs from mine. As he uses the term, someone is biased toward the near iff he performs the act that leads to the greatest discount-adjusted future welfare level.

There is, of course, also room for a third definition of "the bias toward the near" that parallels my definition of "biased (toward the future) attitudes." If someone has this bias, then the degree to which he takes pleasure in the fact that some pleasure is in his future is greater when the pleasure is in the nearer future.

⁷Elizabeth Harman (MS) describes the bias toward the actual (she doesn't call it that) and claims that it is not irrational.

⁸I doubt that it is a necessary truth that getting what we want gives us pleasure; Feldman

preferences. Then I want to have sensory pleasure in my future. If I learn that I will, in fact, experience sensory pleasure in the future, then I (learn that I) am getting what I want. And I take pleasure in this fact. Similarly, since I lack a desire to have experienced pleasure in the past, when I learn that I did experience pleasure in the past, I will get nothing out of the information. I won't take any pleasure in the fact that I experienced some pleasure.

So my biased attitudes are generated by my biased preferences, together with a general connection between desire and pleasure (that has nothing special to do with time). This has consequences for attempts to evaluate the bias toward the future. If we want to know whether it is rational to be biased toward the future, we should ask whether it is rational to have biased preferences. For, given the way that biased attitudes are generated by biased preferences, biased attitudes will be rational just in case biased preferences are.⁹

References

Feldman, Fred (1997). "Two Questions About Pleasure." In *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy*, chapter 7. New York: Cambridge University Press.

— (unpublished). "Pleasure and Desire."

Harman, Elizabeth (MS). "The Mistake in 'I'll be Glad I Did It' Reasoning: The Significance of Future Desires."

Heathwood, Chris (2005). *Desire-Satisfaction Theories of Welfare*. Unpublished Dissertation. Available online at <http://spot.colorado.edu/~heathwoo/DSTW.pdf>.

Kagan, Shelly (1986). "The Present-Aim Theory of Rationality." *Ethics* 96: 746–759.

(unpublished) presents some convincing counterexamples.

⁹Thanks to Chris Heathwood, Fred Feldman, Elizabeth Harman, and Harman's Spring 2008 graduate seminar at Princeton for helpful comments.

Moller, Dan (2002). "Parfit on Pains, Pleasures, and the Time of Their Occurrence." *Canadian Journal of Philosophy* 32: 67–82.

Parfit, Derek (1984). *Reasons and Persons*. New York: Oxford University Press.