Wellbeing and Changing Attitudes Across Time

Krister Bykvist[1,2]

Abstract

The fact that our attitudes change poses well-known challenges for attitude-sensitive well-being theories. Suppose that in the past you favoured your adventurous youthful life more than the quiet and unassuming life you expected to live as an old person; now when you look back you favour your current life more than your youthful past life. Which period of your life is better for you? More generally, how can we find a stable attitude-sensitive standard of wellbeing, if the standard is in part defined in terms of unstable attitudes? In this paper, I introduce an 'attitudinal matrix' framework that will help us clear up the problems posed by changing attitudes across time. In particular, it will help us see what is at stake, which principles that can or cannot be combined, and what might be the best solution. I defend a very plausible candidate constraint on a solution to the challenge of changing attitudes, which I call 'diagonalism'. It is argued that among the three main forms of substantive attitude-sensitive wellbeing theories – the attitude-version, the object-version, and the satisfaction-version – it is the satisfaction-version that can both satisfy diagonalism and provide the best account of temporal and lifetime wellbeing.

1 Introduction

It is a commonplace that our attitudes towards important aspects of our lives can change. Such a change can be prompted by many different events, including choosing a career, getting married, forming a new friendship, becoming a parent, fighting in a war, becoming ill (e.g., Alzheimer's), 'coming out' as homosexual, undergoing gender-modification, being 'cured' of a disability (e.g., blind to seeing or deaf to hearing), starting/stopping taking drugs, being adopted, moving to a new country (with a new culture), exiting one's culture, family, or group of friends, converting to a new religious/atheist/political view, growing old (and conservative).

✉ Krister Bykvist
krister.bykvist@philosophy.su.se

1 Department of Philosophy, Stockholm University, Stockholm, Sweden

2 Institute for Futures Studies, Stockholm, Sweden

Some of these cases involve our *own choices*, e.g., choosing a career, forming a friendship, getting married, and becoming a parent, whereas others involve the *choices of others*, e.g., being adopted, being cured of a handicap, and being forced to leave one's family. Some cases are better seen as things that just *happen* to us, e.g., becoming ill and growing old. What is common to all of these cases is that they typically involve a change in one's fundamental attitudes towards certain important aspects of one's life.

It is well-known that these cases of attitude change pose a challenge for desire-based wellbeing theories, for how are we supposed to assess a person's wellbeing on the basis of her desires, if these desires are in a flux? The challenge has a greater reach than this, however, for it is enough to opt for *attitude-sensitive* wellbeing theory to be affected by it. According to attitude-sensitive wellbeing theories, whether your life (or some part thereof) is good for you, bad for you, or better for you than some other life (or part thereof) depends at least *in part* on the match (or mismatch) between your attitudes and these lives (or their parts). So if the relevant attitudes change, then again we have the question of how to decide which impact the attitudes have on wellbeing. The attitudes in question need not be limited to desires; they can include emotions, intentions, wishes, value seemings, and perhaps even value judgments (about other values than wellbeing to avoid circularity). An attitude-sensitive theory need not be exclusively attitude-based either. Other factors can also affect wellbeing. For example, whether your life is good for you can also depend on whether it is *objectively valuable* in some respect (e.g., aesthetic, artistic, athletic, intellectual, or moral). In fact, I think such a hybrid view has much going for it, but in this paper I am going to put the objective value aspect aside, and just focus on the attitudinal part of an attitude-sensitive wellbeing theory.

It is important to note that this challenge cannot be easily met just by focusing on *ideal* attitudes, the attitudes we would have under certain ideal circumstances, for on the standard accounts of ideal attitudes, these attitudes can change too. Take, for example, the famous account in Harsanyi (1990, 55), according to which ideal attitudes (or 'rational' as he would call them) are the ones you would have, if you were fully informed about relevant empirical facts and reasoned rationally. Such ideal attitudes can change too, since what you would care about if you were rational and more informed in this way depends crucially on your psychological make-up and what kind of person you are. Of course, one could always cook up a definition of ideal attitudes that would make them necessarily stable. For instance, one could define ideal attitudes as the ones you would have, if you cared about a certain unique set of valuable states of affairs and only cared about them to the extent that they are valuable. But then one could wonder why such a theory should merit the label attitude-*sensitive*, since attitudes are an idle wheel in these theories.

Now, attitudes can change in three ways, across *time* (your attitudes at one time differ from those at another), across *worlds* (your attitudes in one world differ from those in another), and across both. In this paper my focus is exclusively on the change across time in one life.[1] I shall first, in Sections 2 and 3, give a more precise statement of what the challenge of changing attitudes consists in. What exactly is at stake here? Then, in Section 4, I shall introduce a framework – the 'attitudinal matrix' framework – that will help us state the challenge clearly and succinctly and assess and compare various solutions thoroughly. In particular, the framework will make it very easy to formulate various conditions on solutions to the challenge, and decide which solutions satisfy which conditions. In Sections 5

---

[1] I have dealt elsewhere with change across worlds, see Bykvist (2010).

and 6, I shall present and defend what I think is a very plausible constraint on a solution to the challenge of changing attitudes, which I will call 'diagonalism'. I shall show what I think are the crucial virtues of this constraint and also contrast it to some other alternative constraints. In Sections 7 and 8, I shall argue that among the three main forms of substantive attitude-sensitive wellbeing theories – the attitude-version, the object-version, and the satisfaction-version – it is only the satisfaction-version that can both satisfy diagonalism and provide the best account of temporal and lifetime wellbeing. Finally, Section 9 concludes the discussion.

2 The Challenge of Changing Attitudes

To put the problem of changing attitudes across time in sharp relief, consider the following case, 'adventurous youth versus quiet old age' (Nagel 1970 1974, 70). In the past you favoured your adventurous youthful life more than the quiet and unassuming life you expected to live as an old person; now when you look back you favour your current life more than your youthful past life. Which period of your life is better for you? More generally, is there a stable standard of wellbeing we can appeal to in these cases, or do we have to accept that the wellbeing value of a life (or part of a life) can change across times? The structure of the case is this:

> In the past, during your youthful life: you favour yours adventurous youthful life more than the quiet and unassuming life you expect to live as an old person;
> In the present, during your quiet old age: you favour your current quiet life more than your adventurous youthful past life.

Which time is better for you? One option is to claim there is no unique answer to this question for comparative wellbeing is *relative to times*:

> In the past, during my youthful life: your adventurous youthful life is better for you than the quiet and unassuming life you expect to live as an old person.
> In the present, during my quiet old age: your current quiet life is better for you than your adventurous youthful past life.

But this would not provide us with a *stable* answer to the question of which time is better for you! This means that there is no stable answer to whether things are becoming better or worse for you, and we can't say whether you are better off or worse off now than in the past. This is a big drawback. In this paper, I shall therefore instead explore non-relative accounts of temporal wellbeing.

3 Attitude-Sensitive Wellbeing Theories

Before we get started with this exploration, we need to get a firmer grip on what attitude-sensitive accounts of wellbeing amount to. As I said earlier, the slogan is that whether your life is good for you, bad for you, or better for you than some other life *depends in part* on the match between your attitudes and these lives. The attitudes in question can be desires, wishes, emotions, preferences, and evaluative responses, but for my discussion there is no need to take a stand on which kind of attitudes count, except that I will assume that they

come with a certain polarity, positive (pro-attitudes), negative (con-attitudes), and neutral (attitudes of neutrality). I will say a bit more about polarity later on.

This dependency between wellbeing and attitudes can be understood in different ways. In this paper, I am interested in *substantive first-order* accounts of attitude-sensitive wellbeing, according to which attitudes are wellbeing-*makers*, or value-for-makers: they explain why something has a certain value for you. This is analogous to substantive first-order normative theories that provide right-makers, properties of actions that explain why they are right.

We can distinguish the first-order wellbeing accounts neatly by what they would say about the (schematic) fact that you favour x and x obtains. According to the *attitude-version*, your *favouring* of x is good for you, because it is satisfied (x obtains). According to the *object-version*, *x* is good for you, because you favour x (assuming no other attitudes are directed at x). Finally, according to the *satisfaction-version*, the *complex state of affairs* that you favour x and x obtains is good for you, because it is an instance of a satisfaction (a combination of a favouring together with the obtaining of its object).[2]

I should add that I take these three theories as possible theories of what has *basic* value for you. Other things, such as broader situations and whole lives, have value for you in virtue of containing things that have basic value for you.

4 The Attitude Matrix Framework

In order to facilitate the discussion I will introduce an attitude matrix framework.[3] I shall assume we can represent the attitudes at different times in a life by a matrix:

	Times	
Attitudes	T1	T2
T1	a	b
T2	c	d

Of course, this is only a very schematic and unrealistic representation of a life, since it consists of only two times, T1 and T2. But this simplicity is a virtue, for it makes it much easier to isolate and handle the specific issues that the change of attitudes across time brings to the fore. All the cases I will discuss could have been given more realistic representations, but that would have made the discussion much more cumbersome.

Here is how you are supposed to interpret the matrix. If you look into a horizontal row, you will find a representation of the person's attitudes (and their degrees), at a certain time, towards different times in her life. So, if you look into the T1-row, you will find numerical representations of the attitudes the person has *at T1* towards her life at T1 and her life at T2. If you look into the T2-row, you will find representations of her attitudes *at T2*, towards T1 and towards T2. So, *a* represents her attitude at T1 towards T1, *b* her attitude at T1

---

[2] For the distinction between the object-version and the satisfaction-version, see Rabinowicz & Österberg (1996), who distinguishes between 'the satisfaction interpretation' and 'the object interpretation' of preferentialism intrinsic value. See also, Persson (1995) p. 64, who makes a similar distinction between what he calls 'the relational view' and 'the object view'. For a more thorough discussion on the normative relevance of this distinction, see Bykvist (1998).

[3] A similar framework for changing attitudes across *worlds* is presented in Bykvist (2010).

towards T2, *c* her attitude at T2 towards T1, and *d* her attitude at T2 towards T2, where *a*, *b*, *c*, and *d* are numbers.

Positive numbers represent favourings (a greater positive number represents a stronger positive attitude), negative numbers disfavourings (a greater negative number represents a stronger negative attitude), and zero, neutral attitudes. A preference, at Ti, for the life in Tj over the life at Tk is represented by a greater number in the entry Ti,Tj than in the entry Ti,Tk. Indifference, at Ti, between Tj and Tk is represented by assigning the same number to both Ti,Tj and Ti,Tk. So, in this simple model, *full comparability* of attitude strength is assumed, both within and across times, (an assumption that can be relaxed in more realistic models). For most of my discussion, I shall only assume that the numbers have *ordinal* relevance (so we can't compare the differences of attitude strengths).

This framework assumes that attitudes come with a polarity: positive (favouring), negative (disfavouring), or neutral (neutral attitude).[4] To *favour* something in the fullest sense is to be positively orientated towards it in one's *experiences*, *motivations*, and *evaluative seemings*: thinking about it is pleasant, being disposed to bring it about, and seeing it as something good (which is not the same as judging it to be good). To *disfavour* something in the fullest sense is to be negatively orientated towards it in the same domains: thinking about it is unpleasant, being disposed to not bring it about or to prevent it from occurring, and seeing it as something bad. Finally, to be completely *neutral* towards something is to be neutrally orientated towards it in the same domains: thinking about it is neither pleasant nor unpleasant, not being disposed to bring it about nor being disposed to not bring it about or prevent its occurrence, and seeing it as something neutral.

The attitudes are supposed to be final ('the attitudes for something for its own sake'), unconditional ('not given some condition') attitudes. Finally, it is also assumed that the attitudes are equally autonomous, rational, and informed, and that, at each time, the attitudes are maximally integrated (no conflicts between the attitudes at that time). This is done in order to be able to focus more specifically on the issues about changing attitudes across time.

All attitudes are supposed to be held at a certain time and in a certain life. In order to avoid taking a stand on the question of how the duration of a certain attitude or its object affects wellbeing, I shall assume that they and their objects all have the same duration. So, in all cases to be discussed T1 has the same duration as T2. Depending on the case, this duration can be equal to the duration of either a moment, seen as a short period – never a durationless instant, however, since it is doubtful we have attitudes towards such instants – or a longer period. I assume that there is no change of attitudes during the relevant period. So, to say that you have a certain attitude at T1 towards T2 (where T1 may be identical to T2, in which case we have a synchronic attitude) is to say that throughout T1 you have this attitude towards T2, and that you have no other attitudes towards T2. A full account would have to decide on the relevant periods to consider when defining momentary or temporal wellbeing, but I am not going to decide on such a partition here.

Finally, when I say you have an attitude towards a time T, I mean that you have an attitude towards *your life at T*, where 'your life' picks out the life that the matrix is supposed to represent. I am not assuming that you need to conceptualize the object of your attitude as 'your life at T'. After all, we rarely, if ever, have attitudes towards what we

---

[4] The following account of the polarity of attitudes is inspired by Hurka (2001, 13–14).

conceptualize as particular times in particular lives. What I mean is rather that you have an attitude towards (a subset of) the propositions that together make up your 'life story' about a certain time in a certain life. This life story gives a complete description of what is going on in your life at that time. In the following, I shall assume that you are able to form an *overall* attitude towards the propositions that together comprise the life story about a certain time in a certain world. It is this overall attitude that will be represented by an attitudinal grade.

Using this framework, we can cash out the case 'adventurous youth versus quiet old age' in the following way (the numbers are not unique, of course, we just need to make sure that (Y,Y) is greater than (Y,O) and (O,O) is greater than (O,Y)):

Case 1

		Times	
Attitudes		Y	O
	Y	4	1
	O	1	2

Filling in the case in this way means that at Y, you favour Y (to degree 4) and favour O less (to degree 1), and at O, you favour O (to degree 2) and favour Y less (to degree 1). We can also see that your Y-self favours Y more than your O-self favours O (degree 4 versus degree 2).

5 Diagonalism

How should we decide whether one time is better for you than another, or vice versa? One straightforward and simple idea, which I find very attractive, is to look at the (principal) diagonal (a, d):

		Times	
Attitudes		T1	T2
	T1	a	b
	T2	c	d

Diagonalism, as we may call it, would then say that

T1 is better for you than T2 iff $a > d$.
T2 is better for you than T1 iff $d > a$.
T1 is equally as good for you as T2 iff $a = d$.

More generally, diagonalism claims that

T is *good* for you iff at T, you *favour* T (and the stronger you favour it, the greater positive value it has for you).

T is *bad* for you iff at T, you *disfavour* T (and the stronger you disfavour it, the greater negative value it has for you).

T is *neutral* for you iff at T, you are *neutral* towards T.[5]

T is *better* for you than T* iff the *value* of T for you is *greater* than the *value* of T* for you.

T is *equally as good* for you as T* iff the *value* of T for you is *identical* to the *value* of T* for you.

What are diagonalism's verdicts on 'adventurous youth versus quiet old age'? Well, that depends on exactly what attitudes you would have towards these periods. Suppose it looks like we depicted it above in Case 1. Then diagonalism would say that both Y and O are good for you, but Y is better for you than O.

One clear advantage of diagonalism is that it provides us with a stable *non-relative* standard of comparative and absolute wellbeing. Whether a time is better for, worse for, or equally as good as another time in your life, whether one time is good, bad, or neutral for you is not relative to times in your life.

Diagonalism has two further important virtues: it captures the so-called *resonance constraint* on wellbeing (under a plausible interpretation of this constraint), and it can deal with changes in your ability to grasp different parts of your life, such as transformative experiences that one cannot understand or grasp until one has had them.[6] Take the resonance constraint first. According to a plausible interpretation of this constraint, a time T (towards which you have an attitude that does not change during T) is good for you only if T 'resonates' with you *at T*. As Bradley (2016, 7) puts it:'In order for something to go well for someone at a time, it must not only resonate with the person at the time – it must also *obtain* at that time'.

Here 'T resonates' means not being alienated from T, caring about T, and finding T (or some parts of it) 'compelling, attractive, engaging' (Railton 2002, 47). This is a plausible interpretation of the constraint, since after all, it would be odd to claim that a time in your life that will leave you completely cold when you lead it still resonates with you.

This means that T 'resonates' with you at T only if you have an T-located and T-directed favouring. For if it is not T-located it does not resonate with you *at T*, and if it is not T-directed, then *T* does not resonate with you. So, we can conclude that T is good for you only if you have an T-located and T-directed favouring. But this is exactly what diagonalism says. So, it captures the resonance constraint.

Diagonalism can also deal with changes in your conceptual abilities to grasp different parts of your life without forcing us to accept incomparability of wellbeing. To see this, consider the following schematic case:

Case 2

Attitudes	Times	
	T1	T2
T1	a	
T2		d

Suppose that at T1 that you are too young to grasp your adult life at T2, and that at T2 you are too old to remember life at T1. So, you have no diachronic attitudes, a fact we represent with two blanks.

This is not a problem for diagonalism, since T1 can still be compared with T2, even though at T1 you have no attitude towards T2, and at T1, no attitude towards T2. We only need to compare the attitude you had at T1 towards T1 with the attitude you have at T2 towards T2. So, suppose, for example, that a is greater than d. Then, according to diagonalism, you were better off at T1 than at T2.

It might look like diagonalism is only compatible with the object-version, since attitudinal values are assigned to the *objects* of attitudes in the matrices above. But this is not so. Diagonalism is in fact compatible with all three versions of attitude-sensitive theories can be made compatible with diagonalism, for they can all accept that T is *good* for you iff at T, you *favour* A. Note that if, at T, you favour T, then T contains:

(a)   Your satisfied *favouring* at T.
(b)   The *object* of your favouring at T, namely T itself.
(c)   The *combination* of your favouring and its object.

This means that all three versions can agree that T is good for you, but for different reasons. The attitude-version will say it is good for you because it contains your satisfied favouring at T, the object-version can say that it is good for you because it contains the object of your favouring at T, and the satisfaction-version that it is good for you because it contains the combination of your favouring at T and its object, T.

Furthermore, since all three versions can be made compatible with diagonalism, they can all accept the resonance constraint. They can all agree on the constraint that *only* T-located and T-directed attitudes are relevant for the value of T for you. I take this to be good news for these versions, since the resonance constraint is intuitively compelling.

The fact that all three version can be made compatible with diagonalism shows that it is better to see it as a *constraint* on a first-order attitude-sensitive theory than such a first order theory in its own right.

6 Too Little Weight to Preferences?

Some might complain that diagonalism is flawed, because it gives too little weight to preferences. For example, it does not comply with

Pairwise Dominance
If both your T-self and T*-self prefer T to T*, then T is better for you than T*.

Here the talk about your T-self (or T*-self) is not supposed to be committed to the metaphysically controversial claim that a person is literally made up of person-like selves. To say that your T-self prefers T to T* is just a more convenient way of saying that, at time T, you prefer T to T*.

To see that diagonalism fails to comply with Pairwise Dominance, consider the following schematic example:

Case 3

	Times	
Attitudes	T1	T2
T1	-2	-4
T2	4	2

Your T1-self disfavours T1, but disfavours T2 even more. Your T2-self favours T2 but favours T1 more. So both your T1-self and your T2-self prefer T1 to T2. But diagonalism would say that T2 is better for you than T1. Here diagonalism seems to be in the right, however, since T1 does not resonate with you at T1, indeed you disfavours it at T1, whereas T2 is something you favour at T2, and thus something that resonates with you at T2.

One might object to this case by claiming that if you knew at T2 about your attitudes at T1, and thus knew that at T1 you *hate* your life at T1, then you would no longer favour, at T2, your life at T1. This response will work for some cases. It will work for those cases in which your attitude is *conditional on its own persistence*: at T, you favour T* only on the condition that at T* you favour T*.[7] But, of course, one's attitudes towards a time might be based on *personal ideals*, and it is a characteristic (if not defining) feature of ideals that they are not conditional on their own persistence. For example, my desire now to be an honest and healthy person in the future is not conditional on my desiring it then. I want now that I am honest and healthy even in the future scenario in which I have become dishonest and lazy and hate honesty and health.

Note also the Pairwise Dominance leads to *cyclical* wellbeing as in this case:

Case 4

	Times		
Attitudes	T1	T2	T3
T1	2	1	3
T2	3	2	1
T3	1	3	2

According to Pairwise Dominance, T1 is better for you than T2, since both your T1-self and your T2-self prefer T1 to T2. T2 is better for you than T3, since both your T2-self and your T3-self prefer T2 to T3. But T3 is better for you than T1, since both your T3-self and your T1-self prefer T3 to T1. So, we have a cyclical wellbeing ordering.

Note that Diagonalism does not even comply with

Comparative Endorsement
If T is better for you than T*, then either your T-self prefers T to T* or your T*-self prefers T to T*

But that seems fine, in light of the intuitive verdicts about Case 3, but also in this case:

---

[7] This kind of conditionality is discussed in Parfit (1992, 151).

Case 5

	Times	
Attitudes	T1	T2
T1	20	20
T2	-20	-20

Your T1-self favours T1 and T2 equally strongly, whereas your T2-self disfavours T2 and T1 equally strongly. It seems clear that T1 is better for you than T2, since your T1-self strongly favours T1 and your T2-self strongly disfavours T2. But neither your T1-self, nor your T2-self prefers T1 to T2. So, Comparative Endorsement is not satisfied.

7 The Relevance of Diachronic Attitudes

Two features of diagonalism need to be highlighted:

(a) diachronic attitudes *at* T (directed towards different times) cannot affect the wellbeing at T.
(b) diachronic attitudes (located at times different from T) *towards* T cannot affect the wellbeing at T.

(a) and (b) together entails that only *synchronic* attitudes (i.e., attitudes at T towards T) count towards the wellbeing at T. No one would deny that synchronic attitudes at T should count towards the wellbeing at T, but should only synchronic attitudes count towards *lifetime* wellbeing? Does *diachronic* attitudes have no effect at all on lifetime wellbeing? Answering 'yes' to these questions would leave us with *synchronism* about life time wellbeing, according to which only the synchronic attitudes at a time determine the wellbeing at that time and then this kind of temporal wellbeing determines lifetime wellbeing. But synchronism is implausible, since it gives no weight to our attitudes for how things should *unfold in time*. (Of course, I do think synchronism about temporal wellbeing is plausible, since this is what diagonalism boils down to.) But we do not just want things to happen at specific times; we also want them to happen in a certain order.[8] For instance, I want to work hard before I receive some gratitude, I prefer an intimate relationship that starts poorly and ends well to one that starts well and ends poorly, and I want to have of period of moderate pleasure (or perhaps even some moderate pain) before I go through some heights of ecstasy. Reversing the order of events may make them look less attractive. The synchronist's *snapshot view*, i.e., the exclusive focus on what happens at particular moments of time, prevents her from taking into account attitudes about temporal wholes. But clearly this would exclude too much since almost all of our projects, commitments, and plans concern getting things done at the right time and in the right order.

That it is important to take into account diachronic attitudes is also brought out by the following cases. In the first case (Case 6), you are raised in adversity, which you deplore, but wish for a fulfilling future, which you in fact will have and will favour (Velleman 1993, 340). The case looks like this:

---

[8] For an order-sensitive account, see Dorsey (2015).

Case 6

	Times	
Attitudes	T1	T2
T1	-1	2
T2	-1	2

Here diagonalism would say that T2 is better for you than T1. The *attitude-version* could count the diachronic attitude at T1 as part of your wellbeing at T1, since it will be satisfied. But it seems implausible to say that your wish for a fulfilling future, which will be satisfied, mitigates the badness of your childhood. After all, your life is not fulfilling yet. But it also seems clear that the fact that your diachronic attitude at T1 for T2 is satisfied should count for something. Finally getting the fulfilling life you always wanted should count towards your lifetime wellbeing.

In the next case (Case 7), 'The manuscript', you tragically fall into an attitude-less coma right after you have finished your opus magnum.

Case 7

	Times	
Attitudes	T1	T2
T1	1	2
T2		

At T1, you favour T2 because you favour the manuscript being sent off and being published (a diachronic attitude for an attitude-less time). At T2, you are in coma and your manuscript is sent off and is published. Here diagonalism would say that T1 cannot be compared with T2.[9]

However, the *object-version* could say that T2 is better for you than T1, since T2 is the *object* of the attitude at T1. Is this the right thing to say? In the next section, I will argue that it is in fact the wrong thing to say. But, on the other hand, this example shows that it would also be wrong to dismiss the diachronic attitude at T1 all together. It is clear that the diachronic attitude should add something to your wellbeing. Satisfying one's lifetime project should count towards one's lifetime wellbeing, even if one falls into a coma right before completion.

In the last case (Case 8), 'Redeem the past', an unhappy artist had it all but could not endorse her life then because of depression (McKerlie 2012, 170). However, later when she is older she looks back and appreciates her past achievements.

Case 8

	Times	
Attitudes	T1	T2
T1	-1	1
T2	1	1

---

[9] If existing without having any attitudes is a way of having neutral wellbeing, then diagonalism would say that T1 is better than T2.

Does the favouring at T2 towards T1 make T1 less bad? Diagonalism would say no. However, the *object-version* could say this, since the *object* of the favouring at T2 is located at T1.[10] The attitude- and satisfaction-versions cannot say this, since the favouring or the satisfaction is not located at T1. But should we say this? In the next section, I will argue that we should not say this. But, on the other hand, this example makes it clear that the diachronic attitude should add something to your wellbeing. Seeing your past achievements in a better light should count towards your lifetime wellbeing.

8 Diachronic Attitudes as Determinants of Lifetime Wellbeing, Not Temporal Wellbeing

So, it is clear that both synchronic and diachronic attitudes should count towards lifetime wellbeing. One way to count diachronic attitudes is to count them as affecting your temporal wellbeing and thereby affecting your lifetime wellbeing. The attitude-version and the object-version seem to have no other way of counting diachronic attitudes. The attitude-version has to say that if the diachronic attitude at T counts, then this affects the wellbeing at T, since what has value for you – the attitude – is located at T. The *object-version* has to say that if the diachronic attitude towards T counts, then this affects the wellbeing at T, since what has value for you – the object of your attitude – is located at T. But these are problematic implications.[11]

First of all, by letting diachronic attitudes determine temporal wellbeing in this way, the attitude-version and the object-version violate the intuitive intrinsicality constraint on temporal wellbeing, which says that your wellbeing at a time T is a fact *intrinsic* to T and not dependent on your relations to other times.[12] The attitude-version would have to say that your diachronic attitude at T towards T*, if it will be satisfied, adds to your wellbeing at T even though its satisfaction is not a fact intrinsic to T, since it involves a relation between your attitude at T and your life at T*.[13] Similarly, the object-version would have to say that your diachronic attitude at T towards

---

[10] Many philosophers seem attracted to retroactive benefits, including Bigelow et al. (1990, 119-40), Feinberg (1984, 79-85), and McKerlie (2012, 170).

[11] Without talking about the object-version or the attitude-version, Baber (2010) defends what Bradley (2016) calls 'the time of the object view', which, like the object-version, states that you are benefited by the time at which the object of your attitude obtains. Bruckner (2013) and Dorsey (2013) defend what Bradley (2016) call 'the time of the attitude view', which, like the attitude-version, states that you are benefited at the time at which you hold the attitude. See, Bradley (2016) for criticisms of Baber (2010), Bruckner (2013), and Dorsey (2013).

[12] This constraint is discussed and defended in Velleman (1993, 339) and Bradley (2009). Bradley (2016) also makes use of this principle to criticize 'the time of the object view' and 'the time of the attitude view', described in footnote 11.

[13] One might object to the intrinsicality constraint by referring to cases in which you are now pleased that the past turned out in a certain way. It does not seem strange to say that this past-looking emotion adds to your current wellbeing. I agree, but it is important to ask what it is about this emotion that adds to your current wellbeing. If our intuition trades on the fact that you now have an *experience of pleasure*, then this is not a counterexample to the intrinsicality constraint, since the experience occurs now. It is instead a counterexample to an attitude-sensitive view that does not count pleasant experiences as wellbeing-makers, which include all the theories I consider in this paper. In other words, if you want to count this experience of pleasure as a wellbeing-maker, you have to go beyond the monistic attitude-sensitive wellbeing theories I am discussing in this paper (unless you define a pleasure as an experience that you have a positive attitude towards). On the other hand, even a monistic attitude-sensitive theory can tell a story, perhaps a bit contrived, about why being pleased about the past adds to our wellbeing without violating the intrinsicality constraint, for it can claim that if you now care about feeling this pleasure, then satisfying this synchronic pleasure-oriented attitude adds to your current wellbeing. Whether we should go beyond monism or instead tell this story is not something I want to take stand on here. Thanks to Monika Betzler for pressing me on this.

T* adds to your wellbeing at T*, even though the fact that T* is the object of the diachronic attitude at T is not a fact intrinsic to T*, since it involves a relation between T* and your diachronic attitude at T.

However, I do not want to hang to much on the intrinsicality constraint, since it is not endorsed by everyone (it seems to rule out that memory or knowledge of the past can make my life better now).[14] What is more important, and also independent of the standing of this constraint, is that both the attitude-version and the object-version run into further problems about temporal wellbeing, if we make the plausible assumption that a favouring to degree n and a disfavouring to a degree –n exactly balance each other in strength so that the result is something neutral for you.[15] Take the attitude-version first and suppose your life looks like this:

Case 9

		Times	
Attitudes		T1	T2
	T1	-2	2
	T2	-4	4

The attitude-version would have to say that, at both T1 and T2, your life is neutral for you, since your synchronic attitude at each time is exactly balanced by your diachronic attitude at that time.[16] So, there is *no improvement* over time in your life, despite the fact that at T1 you disfavour T1 and at T2 you favour T2.

The object-version has an analogous problem. Suppose your life looks like this.

Case 10

		Times	
Attitudes		T1	T2
	T1	-2	-4
	T2	2	4

The object-version would have to say that both at T1 and T2 your life is neutral for you, since the synchronic attitude at each time is exactly balanced by the diachronic attitude towards this time. Again, there is *no improvement* over time, despite the fact that at T1 you disfavour T1 and at T2 you favour T2.

In contrast, the satisfaction-version has no problem complying with the intrinsicality constraint, since it can say that only the satisfaction of synchronic attitudes at T count towards your wellbeing at T, and these satisfactions are facts intrinsic to T. This also means that, in the cases above, the satisfaction-version can say that there is an improvement in wellbeing, since at T1 you disfavour T1 and at T2 you favour T2.

How can the satisfaction-version count diachronic attitudes without seeing them as determinants of temporal wellbeing? It can do so by saying that diachronic attitudes

---

[14] Thanks to Maria Lasonen-Aarnio for this example.

[15] Here I assume more than a merely ordinal significance of attitudinal grades. Note also that for my purposes it is sufficient to assume that there are grades n and m, *where n need not be identical to m*, such that a favouring to degree n and a disfavouring to degree m will cancel each others out so that the result is something that is neutral for you. I assume that n is identical to m in the Case 9 just for simplicity.

[16] Velleman (1993, 398) suggests something similar.

can have an effect on *lifetime* wellbeing without having an effect on *temporal* wellbeing. The satisfaction of your diachronic attitudes provide a better overall structure (narrative) – a temporally holistic good that makes the life better for you without making it better for your at any particular time.

This holistic good is either *atemporal* or *temporally scattered*. According to the first option, your life has a better overall structure by including satisfied diachronic attitudes, but at no specific time does it have this overall structure.[17] According to the other option, the satisfaction of a diachronic attitude is a holistic good that is located at the temporally scattered *fusion* of the time of the diachronic attitude and the time of the object of the attitude. Both options enables the satisfaction-version to say that a diachronic attitude can make a difference to lifetime wellbeing without making a difference to temporal wellbeing.

Of course, much more has to be said about the relevance of diachronic attitudes for lifetime wellbeing. For example, we need to decide exactly how much weight to give to diachronic attitudes compared to synchronic ones. This is a difficult question and the answer to it may depend on many factors. Plausible candidates include: (a) the duration of the diachronic attitudes, (b) the genesis of the attitudes, (c) the psychological connectedness between the subject at the time of the diachronic attitude and subject at the time of the synchronic one, and (d) the objective values of the objects of attitudes.[18]

9 Concluding Remarks

I have argued that diagonalism is an important constraint on temporal wellbeing and that the satisfaction-version of attitude-sensitive wellbeing theories is superior to the attitude-version and the object-version in the way temporal and lifetime wellbeing are characterized. Both the satisfaction of synchronic attitudes and the satisfaction of diachronic ones determine lifetime wellbeing. But whereas the satisfaction of synchronic attitudes determines temporal wellbeing, the satisfaction of diachronic attitudes provide holistic goods that cannot be located at particular times. Of course, much more has to be said about exactly how diachronic attitudes determine lifetime wellbeing. But I hope that, even if you have doubts about my conclusions, you at least appreciate the matrix framework I have presented, since it provides a very simple model of changing attitudes that, partly because of its simplicity, makes it much easier to compare the advantages and disadvantages of different attitude-sensitive theories.[19]

---

[17] This option is defended in Bradley (2016).

[18] For further discussion of such factors, see Bykvist (2003) and Pettigrew (2020). For an overview of how diachronic and synchronic wellbeing can determine lifetime wellbeing, see Kauppinen (2015).

[19] I would like to thank the audiences at the workshop 'Time and the Good Life' at Bochum University, 20th of May, 2021, the workshop 'Explaining value change' at Delft University, May 6th, 2021, and the departmental seminar at the Department of Philosophy, University of Helsinki, 29th of April, 2022. I am especially grateful for comments from Monika Betzler, Eva Weber-Guskar, Holmer Steinfath, Dan Haybron, Valerie Tiberius, Antti Kauppinen, Maria Lasonen-Aarnio, Jaakko Hirvelä, and two anonymous referees.

Data Availability n/a.

Code Availability n/a.

Declarations

Conflicts of interest/Competing interests n/a.

References

Baber HE (2010) Ex Ante desire and post Hoc satisfaction. In: Campbell J, O'Rourke M, Silverstein H (eds) Time and Identity. MIT Press, Cambridge, pp 249–267

Bigelow J, Campbell J, Pargetter R (1990) Death and wellbeing. Pac Philos Q 71:119–140

Bradley B (2009) Well-being and death. Oxford University Press, Oxford

Bradley B (2016) Wellbeing at a time. Philos Exch 45(1):1–12

Bruckner D (2013) Present desire satisfaction and past well-being. Australas J Philos 91:15–29

Bykvist K (2003) The moral relevance of past preferences. In: Dyke H (ed) Time and ethics: Essays at the intersection. Kluwer Academic Publishers, Dordrecht, pp 115–136

Bykvist K (2010) Can unstable preferences provide a stable standard of well-being?'. Econ Philos 26(1):1–26

Bykvist K (1998) Changing Preferences. A Study in Preferentialism, Uppsala University, Repro HSC

Dorsey D (2013) Desire-satisfaction and Welfare as temporal. Ethical Theory Moral Pract 16:151–171

Dorsey D (2015) The significance of a life's shape. Ethics 125:303–330

Feinberg J (1984) Harm to others. Oxford University Press, Oxford

Harsanyi J (1990) Morality and the theory of rational behaviour. In: Sen A, Williams B (eds) Utilitarianism and beyond. Cambridge University Press, Cambridge

Hurka T (2001) Virtue, vice, and value. Oxford University Press, Oxford

Kauppinen A (2015) The narrative calculus, vol 5. Oxford Studies in Normative Ethics, Oxford, pp 196–220

McKerlie D (2012) Justice between the young and the old. Oxford University Press, Oxford

Nagel T (1970) The possibility of altruism. Princeton University Press, Princeton

Parfit D (1992) Reasons and persons. Clarendon Press, Oxford

Paul L (2014) Transformative Experience. Oxford University Press, Oxford

Persson I (1995) Peter Singer on why persons are irreplaceable. Utilitas 7:55–66

Pettigrew R (2020) Choosing for changing selves. Oxford University Press, Oxford

Rabinowicz W, Österberg J (1996) Value based on preferences. On two interpretations of Preference Utilitarianism. Econ Philos 12:1–27

Railton P (2002) Facts and Values. In his Facts, Values, and Norms: Essays Toward a Morality of Consequence. Cambridge University Press, 43–68

Velleman D (1993) Well-being and time. In: Fischer M (ed) The metaphysics of death. Stanford University Press, Stanford, pp 327–357