




ARTICLE

Well-being and the Problem of Unstable Desires

Atus Mariqueo-Russell 

Department of Philosophy, University of Southampton, Southampton, UK
Email: atusmariqueorussell@gmail.com

Abstract

This paper considers a new problem for desire theories of well-being. The problem claims that these theories are implausible because they misvalue the effects of fleeting desires, long-standing desires, and fluctuations in desire strength on well-being. I begin by investigating a version of the desire theory of well-being, simple concurrentism, that fails to capture intuitions in these cases. I then argue that desire theories of well-being that are suitably stability-adjusted can avoid this problem. These theories claim that the average strength of a desire, and the length of time that it is held, both influence the extent to which its fulfilment or frustration affects well-being. I end by considering whether value-fulfilment theories of well-being have a more attractive response to this problem. I find that these theories have significant downsides that make them unappealing alternatives.

Introduction

Desire theories of well-being claim that a subject's well-being is entirely determined by the fulfilment and frustration of their desires. According to one version of this theory, simple concurrentism, the strength of a desire at the time of its fulfilment or frustration determines the extent of its effects on well-being. The problem of unstable desires emerges when we observe that fleeting desires can be strong, long-standing desires can be weak, and desires in general fluctuate in strength. Consequently, simple concurrentism implies that you benefit more from the fulfilment of fleeting desires than is intuitive; that you benefit less from the fulfilment of long-standing desires than is intuitive; and, more generally, that the degree to which you benefit from the fulfilment of a desire varies, implausibly, with fluctuations in its strength. This view has symmetrical implications about the extent of the harms incurred by frustrating desires. For these reasons, I argue that we should reject this view. I then introduce an alternative version of the theory that avoids this problem. Stability-adjusted desire theories of well-being claim that the average strength of a desire, and the length of time that it is held, both influence the extent to which its fulfilment or frustration affects well-being. I find that, rather than undermining desire theories of well-being, reflection on the problem of unstable desires illuminates a way in which these theories can better capture common intuitions about well-being. Consequently, insofar as one finds desire theories

of well-being compelling, stability-adjusted views are an improvement on how these theories are usually constructed.

This paper has the following structure: section 1 outlines desire theories of well-being; section 2 outlines simple concurrentism and the problem of unstable desires; section 3 argues that simple concurrentism cannot avoid this problem; section 4 outlines stability-adjusted desire theories of well-being and defends these views from two objections; section 5 considers value-fulfilment theories of well-being and argues that these views are unattractive solutions to the problem of unstable desires.

1. *Desire theories of well-being*

Well-being is the value that determines how non-instrumentally well a life goes for the subject living it. Desire theories of well-being (henceforth desire theories) are a family of views about what comprises well-being. The following two features constitute a minimally plausible version of the view:

1. Only the fulfilment of a subject's desires non-instrumentally increases their well-being, while only the frustration of their desires non-instrumentally decreases their well-being.
2. The extent to which a subject is made better or worse off by a desire fulfilment or frustration is proportional to the strength of their fulfilled or frustrated desire.

We should note at the outset that the second feature of the proposed minimally plausible desire theory underspecifies at which time during a desire's lifespan its strength determines the extent of its effects on well-being. Consequently, this theory cannot be operationalised until we further specify the second feature to address this deficit. It is possible to quibble that calling this view 'minimally plausible' is a misnomer given that the view itself is unusable without further elaboration. Nevertheless, I favour this terminology because plausible variations of the second feature are likely to be compatible with the second feature as it stands. This is because they are likely to involve further specification of how desire strength should be integrated into the framework. As we will see, the two substantive versions of the desire theory considered in this paper disagree about how to specify this second feature. It is to avoid pre-empting this discussion that I leave this particular issue unsettled here.

The proposed minimally plausible desire theory is vague about whether all desires count equally, whether some count differently, or whether some are excluded from affecting well-being altogether.¹ It also sidesteps debates about how desires affect well-being when their objects occur after the desire has been lost.² This is a strategic necessity to facilitate a focussed discussion of the problem of unstable desires. Nevertheless, the

¹For example, some writers argue that desire theories should exclude instrumental desires from non-instrumentally affecting well-being (Heathwood 2005: 489; Sarch 2013: 223). I remain neutral on this issue because of the difficulty in distinguishing between instrumental and intrinsic desires (Parfit 1984: 117), because the restriction would be misleading if it turns out that instrumental desires do not exist (Murphy 1999: 252–56), and because some intuitions suggest that instrumental desires can be non-instrumentally relevant to well-being (Heathwood 2019: 669).

²For instance, some writers argue that desires do not need to exist at the same time as their objects in order to improve well-being (Bruckner 2013; Dorsey 2013; Vorobej 1998). The view discussed in the bulk of this paper, simple concurrentism, rejects this claim. The theory that I argue for in section 4 is compatible with either position.

findings of this paper can be applied to desire theories that take different positions on these debates. I turn now to consider the problem of unstable desires.

2. *The problem of unstable desires*

Fleeting desires are sometimes strong. This is often the case when they coexist with strong emotions. For instance, when experiencing road rage, we may have a strong desire to shout at other motorists; or, when elated, we may have a strong desire to give away significant amounts of money. Despite the strength of these desires, they are quickly lost. Other desires persist over time but are only moderately held. Take, for instance, the desire to finish writing a book or the desire to perfect a second language. Because these desires are held for a long time, we should expect their strength to fluctuate over time.

Some versions of the desire theory entail that fleeting desires, long-standing desires, and fluctuations in desire strength have counterintuitive effects on well-being. This is the case with simple concurrentism. This view accepts the first feature of the minimally plausible desire theory. However, it specifies the second in the following way:

- 2*. The extent to which a subject is made better or worse off by a desire fulfilment or frustration is proportional to the strength of their desire at the time of its fulfilment or frustration.

Whereas the proposed minimally plausible desire theory underspecifies how the strength of desire affects well-being, simple concurrentism stipulates that it is the strength of desire at the time of its fulfilment or frustration that determines the extent of these effects.

I focus on simple concurrentism for three reasons. Firstly, concurrentist views are influential in the literature (Hare 1981: 103; Heathwood 2005: 490).³ Secondly, whereas desire theories often underspecify how the strength of a desire affects a subject's well-being, simple concurrentism explains this relationship clearly. This clarity serves as a fruitful basis for discussion. Thirdly, framing this discussion around simple concurrentism serves as a useful dialectical device to highlight how desire theories can avoid the problem of unstable desires. As we will see in section 4, findings from this discussion apply to desire theories that reject the claim that desires must exist at the same time as their objects in order to improve well-being.

There are at least five reasons why simple concurrentism is attractive. Firstly, it is conceptually parsimonious. All things being equal, simple theories of well-being are better placed to avoid accusations of arbitrariness than more complex competitors. Secondly, this view captures the resonance requirement (Rosati 1996). This is the intuition that our own well-being must be something that we ordinarily find compelling and attractive, or at least would do if we were more rational and well-informed (Railton 1986: 9). Thirdly, this view might avoid the problem of Dead Sea apples. This term

³The term 'concurrentism' has been used to address several distinct questions in the literature. Here are four questions that can have concurrentist answers: How must a desire be related to its object to count as fulfilled? How must a desire be temporally related to a subject in order to affect their well-being? At what time does the fulfilment or frustration of a subject's desire affect their well-being? How does the strength of a subject's fulfilled or frustrated desire determine the extent of its effects on their well-being? This paper is concerned with this last question.

describes situations where we desire something, get it, and find ourselves disappointed and devoid of feelings of satisfaction (Sidgwick 1907: 110). Dead Sea apples are alleged counterexamples to the thesis that fulfilling desires always improves well-being. Simple concurrentism can claim that, in these cases, the desire ceases immediately prior to its fulfilment and therefore well-being is not improved (Heathwood 2005: 493).⁴ Fourthly, this view can avoid the arguably counterintuitive conclusion that fulfilling lost desires non-instrumentally improves well-being (Bykvist 2003: 116). Finally, this theory can postulate a universal standard of well-being applicable to all subjects capable of well-being. Consequently, its explanatory value is greater than theories that simply purport to explain human well-being. Moreover, rival theories that require species-specific standards of well-being may struggle to explain how to non-arbitrarily identify which standards apply to which species (Lin 2018: 324). While these attractions are not individually exclusive to simple concurrentism, or even to desire theories more generally, they combine to explain why this version of the desire theory warrants extended discussion.

If simple concurrentism is correct, then moderately held long-standing desires matter less to well-being than strongly held fleeting desires. For example, this view entails that the fulfilment of the strong fleeting desire to shout at another motorist improves well-being more than that of the moderate long-standing desire to finish writing a book. Moreover, on this view, fluctuations in desire strength have counterintuitively large effects on well-being. For instance, if the strength of a desire to finish writing a book decreases in the moments preceding its fulfilment, then its fulfilment improves well-being substantially less than had it been fulfilled mere moments beforehand. These are troubling conclusions.

Part of the appeal of desire theories is that they tie well-being to the subjective attitudes that we most closely identify with. This seems to be most true of our long-standing desires. However, the long-standing desires that we tend to consider central to our identity are often prone to fluctuations in strength. They are therefore systematically misvalued by simple concurrentism. This problem is compounded when we consider that we can feel alienated from desires that conflict with our values (Hubin 2003). Alienation seems to be more commonly experienced in relation to fleeting desires. Take, for instance, the strong but short-lived desires of a recovering addict to return to drug use or gambling (Lewis 2000: 70). Because simple concurrentism overvalues the effects that fulfilling and frustrating fleeting desires have on well-being, it places too much importance on those subjective attitudes that we consider peripheral or contrary to our identity.

These concerns form what can be called ‘the problem of unstable desires’. It can be summarised as follows:

- P1:** Viable theories of well-being must intuitively value the effects of fleeting desires, long-standing desires, and fluctuations in desire strength on well-being.
- P2:** Simple concurrentism does not intuitively value the effects of fleeting desires, long-standing desires, and fluctuations in desire strength on well-being.
- C:** Therefore, simple concurrentism is not a viable theory of well-being.

⁴William Lauinger argues that this approach fails because there remain cases where we are desiring, getting, but not benefitting from the object of our desire (2011: 311). If he is right, then one of the attractions of simple concurrentism is revealed as illusory.

This argument is valid. Moreover, P1 is compelling. In the next section, I argue that simple concurrentism cannot avoid P2. Consequently, I find that we ought to reject this view.

3. Simple concurrentism

Common intuitions suggest that the fulfilment and frustration of long-standing desires generally affects well-being more than that of fleeting desires. Proponents of simple concurrentism can attempt to capture this intuition by appealing to a distinction between phenomenological and overall strength of desire. According to desire theories, it is the overall strength of the desire, rather than that of its phenomenology, that determines the extent of its effects on well-being (Griffin 1986: 14–15). If the two are distinct, then proponents of the view can make the following three claims:

1. While fleeting desires may sometimes feel strong, they are normally weakly held.
2. While long-standing desires may sometimes feel weak, they are normally strongly held.
3. While we may sometimes feel fluctuations in desire strength, desires are normally relatively stable.

This approach distinguishes between the overall strength of desire and the strength of its phenomenology. Although this distinction might initially seem *ad hoc*, there are independent reasons to think that it is necessary to accurately explain our mental lives. Call the view that the overall strength of desire is proportional to the strength of its phenomenology ‘proportionalism about desire and phenomenology’, or just ‘proportionalism’ for short.⁵

Proportionalism is not an attractive view. Rejecting it seems necessary to explain a range of psychological phenomena. A striking case occurs when we forget desires (Gregory 2021: 34–6). For example, we may forget our desire to pay off a credit card before incurring fines. Forgotten desires have no phenomenology. Consequently, proportionalism must claim that these desires have been lost. However, it is more intuitive to think of them as simply non-occurrent. Another case is weakness of will. This sometimes seems to happen when our strongest desires fail to produce strong phenomenological effects (Gregory 2021: 34–6). Take, for instance, the abstinent smoker capitulating to the fleeting desire for a cigarette. In this case, it appears that the phenomenology of the desire for the cigarette amplifies motivation and overwhelms the stronger long-standing desire to maintain abstinence. Unusually strongly felt desires also contribute to this picture (Raibley 2010: 598–99). Consider again the desires that emerge during road rage. While many people strongly feel these desires, far fewer would reflectively endorse them as having been strongly held.

Most damagingly, proportionalism entails that we infrequently desire to retain the things that we already have.⁶ These desires often lack a phenomenology.

⁵Mark Schroeder uses the term ‘proportionalism’ to describe the view that the strength of our reasons is proportional to the strength of our desires (2007: 164–70). I am appropriating that term and using it for the claim that the phenomenological strength of a desire is proportional to its overall strength. Elsewhere I discuss a different type of proportionalism that claims that the motivational strength of a desire is proportional to its overall strength (Mariqueo-Russell 2023: §5). There are some parallels between these discussions.

⁶Some readers may be concerned that this argument is incompatible with the death of desire principle. This is the claim that desires always cease to exist once we perceive their fulfilment (Pineda-Oliva 2021:

Nevertheless, it is counterintuitive to claim that we have no such desires. Take, for instance, the desire to remain employed. While its phenomenology can be elicited through prompting, it is usually dormant until we focus our attention or perceive that our employment is threatened. An understanding of desire that identifies overall strength of desire with strength of its phenomenology is unable to explain our strong but phenomenologically dormant desires to retain the things that we have. Consequently, we have independent reasons to reject proportionalism.

However, this approach cannot completely defuse the problem of unstable desires. There is nothing in this picture that disbars fluctuations in desire strength from having counterintuitively large effects on well-being. To prevent this from happening, we would need to postulate a degree of desire stability that far exceeds common intuitions. There appears to be no independent reason to hold such a view. Moreover, rejecting proportionalism does not prevent the fulfilment and frustration of fleeting desires from having counterintuitively large (and long-standing desires from having counterintuitively small) effects on well-being. This argument can only plausibly claim that such occurrences are less frequent than our phenomenology suggests. Therefore, rejecting proportionalism only succeeds in partially diminishing the force of the problem. It fails to solve it.

Even if we are content to live with a diminished problem of unstable desires, there is an additional issue that simple concurrentism faces. Chris Heathwood writes of this:

What if the intensity of the desire changes over the time that it is being concurrently satisfied? We can avoid this by defining desires as things that occur at instants (or at very brief intervals of time). At each brief interval at which some concurrent desire satisfaction occurs, there is just one intensity. We can say that the intensity of a desire has ‘changed’ when a person has a desire of some intensity for some proposition at some brief interval and then, at the next brief interval, has a desire of a different intensity for the same proposition (Heathwood 2005: 490n5).

Here Heathwood addresses the question of how simple concurrentism ought to account for fluctuations in desire strength that happen during the time of fulfilment and frustration. Simple concurrentism claims that the extent to which well-being is affected is determined by the strength of desire at the time of its fulfilment or frustration. However, it is unclear how the view handles cases where there are multiple levels of desire strength during that time. For instance, consider the desire to ride a rollercoaster. This may be strong at the start of the ride, weak during a particularly steep incline, and moderate in the moments preceding the ride’s end. This desire has three different levels of strength during the time of its fulfilment. Simple concurrentism needs to have something to say about such cases.

Heathwood responds to this problem by claiming that there is a single level of desire strength that covers instants or brief intervals of time. If this is right, then examples like the rollercoaster ride are composed of a series of consecutive desire fulfilments about the same thing, rather than fluctuations in desire strength during the time of fulfilment. Call this the desire-at-instant view.

247). However, my objection is based on proportionalism’s failure to recognise our desires to *retain* the things that we already have. These are future-directed desires. Consequently, the objection stands even if the death of desire principle is correct.

The desire-at-instant view can explain how apparent fluctuations in desire strength during the time of fulfilment and frustration affect well-being. However, this approach requires a revision of common intuitions about desires. Our phenomenology and everyday language both point towards desires existing for longer periods of time. Consider your long-term desires to remain in a relationship or employment. These are intuitively better captured by a theory that allows desires to exist for more than a brief interval or instant. Of course, these intuitions may be worth revising if there are independent reasons for accepting the desire-at-instant view. However, to my knowledge, no such reasons exist.

Moreover, this view significantly overvalues the well-being effects of desires for processes and undervalues those of desires for outcomes. Examples of desires for processes might include the desire to play a game of football, watch a film, or run a half marathon. Conversely, examples of desires for outcomes might include the desire to finish writing a book, reach the summit of Mt. Everest, or have nations sign on to an international emissions reduction treaty. Of course, it is possible that these latter examples could also involve desires for processes as well. However, let us set aside that possibility for illustrative purposes. It is true that there are at least some desires that are entirely concerned with processes and at least some others that are entirely concerned with outcomes.

Assume that my examples of desires for outcomes exist alongside their fulfilment for only an instant. The moment the last word is penned, the summit is reached, or the treaty is signed, is the only time that the desire exists alongside its fulfilment. Conversely, imagine that the desires for the processes listed exist for far longer alongside their fulfilment. Assume that the instants that Heathwood has in mind are one second long and that the processes that I have listed last for ninety minutes. That means that these processes involve 5,400 consecutive desire fulfilments. Conversely, the outcomes listed involve only a single desire fulfilment. Consequently, these desires for outcomes would need to be far stronger than the desires for processes to have the same effects on well-being. Clearly, this is an unacceptable result. Desires for processes do not matter overwhelmingly more to well-being than desires for outcomes. Heathwood's view fails to capture this intuition.

A possible response to this problem involves claiming that the benefits of fulfilling desires for outcomes primarily derive from our backwards-facing desires about them. For instance, while fulfilling the desire to reach the summit of Mt. Everest only marginally increases well-being, the fulfilment of backwards-facing desires about this accomplishment can continue to improve well-being long after the event. If this is right, then the gap between our valuations of outcomes relative to processes may narrow. Nevertheless, this approach is unviable in at least some cases. Consider the following counterfactual: The world ends after you fulfil your desire to play a ninety-minute football game, or the world ends after you fulfil your desire to reach the summit of Mt. Everest. In this case, it is not possible to appeal to backwards-facing desires to explain why the well-being effects of the former do not drastically outweigh that of the latter on the view being considered. Even in less dramatic scenarios, this approach does not work. This is especially true of long-standing desires fulfilled later in life. In these cases, we do not have much time remaining to have backwards-facing desires about these outcomes.

While simple concurrentism can partly diminish the force of the problem of unstable desires by rejecting proportionalism, this view struggles to intuitively explain how fluctuations in desire strength during the time of fulfilment or frustration affect

well-being. The only candidate explanation of this in the literature, an appeal to the desire-at-instant view, fails on this account. For these reasons, we should reject simple concurrentism and look for an alternative theory of well-being.

4. *Stability-adjusted desire theories of well-being*

It is possible to construct a version of the desire theory that captures many of the attractions of simple concurrentism and which can solve the problem of unstable desires. This view rejects the idea that the strength of a desire at the time of its fulfilment or frustration determines the extent of its effects on well-being. Instead, it calculates these effects by appealing to two principles. The averaging principle claims that the lifetime average strength of a desire governs the extent of its effects on well-being, while the longevity principle claims that the length of time that a desire is held magnifies these effects.⁷ Call this view ‘the stability-adjusted desire theory’. It accepts the first feature of the proposed minimally plausible desire theory (section 1), but specifies the second in the following way:

- 2***. The extent to which a subject is made better or worse off by a desire fulfilment or frustration is proportional to the lifetime average strength of their fulfilled or frustrated desire and magnified by the length of time that the desire is held.*

The averaging principle prevents fluctuations in desire strength from having counterintuitively large effects on well-being, while the longevity principle entails that long-standing desires generally affect well-being more than fleeting desires. Consequently, this view can value the effects of fleeting desires, long-standing desires, and fluctuations in desire strength on well-being intuitively. This avoids the problem of unstable desires that bedevilled simple concurrentism.

Without further specification, the stability-adjusted desire theory does not entail that desires must exist at the same time as their objects in order to improve well-being. Consequently, this view lacks some of the attractions of simple concurrentism. For instance, it does not entail that fulfilling lost desires does not improve well-being, and it cannot put forward the same solution to the problem of Dead Sea apples as simple concurrentism.⁸ Nevertheless, the stability-adjusted desire theory retains the ability to postulate a universal standard of well-being applicable to all subjects capable

⁷There are precedents for the view that the longevity of a desire magnifies the extent of its effects on well-being (Heathwood 2015: 135). For instance, Krister Bykvist tentatively defends the view that the moral significance of a preference is amplified by its longevity (2003: 128). My view differs from Bykvist’s in two important respects. Firstly, whereas Bykvist claims that desires must exist at the same time as their objects in order to be significant, I remain neutral on this question. Secondly, unlike Bykvist, my view claims that the average lifetime strength of a desire also influences the extent of its significance. Additionally, my paper outlines a distinct motivation for accepting the longevity principle and considers a separate counterargument to it.

⁸There are compensatory advantages to rejecting concurrentism. For instance, some writers argue that lost desires can improve well-being if their objects occur after the desire ceases (Bruckner 2013; Dorsey 2013; Vorobej 1998). One advantage of this approach is that it can postulate posthumous benefits and harms (Pitcher 1984). The debate between concurrentist and non-concurrentist desire theories deserves more attention than is possible here. Consequently, I am content with limiting myself to the claim that both versions are made more plausible by adopting 2**.

of well-being, and can capture the resonance requirement (section 2). If we want this theory to capture the additional attractions of simple concurrentism, then we can modify this view by postulating a third feature:

3. Only desires that exist concurrently with their objects non-instrumentally improve well-being.

The resultant version of the desire theory accepts features 1, 2**, and 3. Call this view 'stability-adjusted concurrentism'. Whereas simple concurrentism entailed 3 through its specification of 2*, stability-adjusted concurrentism must explicitly adopt this feature. This revised theory captures all of the above attractions of simple concurrentism. I remain neutral in this paper between the stability-adjusted desire theory and stability-adjusted concurrentism. The discussion that follows applies equally well to both versions of the desire theory. For simplicity, I will simply refer to stability-adjusted views going forward.

Stability-adjusted views sacrifice some of the conceptual parsimoniousness that made simple concurrentism initially attractive. Notably, they must explain how to balance desire longevity with desire strength. One way of doing so involves identifying units of desire strength and units of temporal length and simply multiplying them together. Furthermore, if we take this approach, then we ought to factor in a diminishing rate of added value for each additional temporal unit prior to multiplication. This prevents the theory from counterintuitively entailing that weakly held long-standing desires matter overwhelmingly more to well-being than much stronger fleeting desires. If we want to fully guard against the possibility of such valuations, then we may want the rate to decrease to zero eventually. Factoring in a diminishing rate of added value for each additional temporal unit means that fulfilling long-standing desires improves well-being progressively more than fulfilling fleeting desires, but not to the extent that fleeting desires become largely irrelevant to well-being. This strikes me as an intuitive result. Nevertheless, postulating a diminishing rate of added value for each additional temporal unit does further compromise the extent to which these theories can remain conceptually simple.

Adopting 2** can solve the problem of unstable desires. Nevertheless, some readers may be concerned that this solution is intolerably arbitrary. Without independent reason for this way of specifying 2, then the adoption of 2** could be perceived as an arbitrary manoeuvre solely made to evade a narrow set of counterexamples. However, I do not think that the charge of arbitrariness is justified. There are at least two reasons for this. Firstly, the way that desire theories are usually constructed underspecifies how exactly strength of desire affects well-being. These theories need to answer this question. Simple concurrentism is no less arbitrary than stability-adjusted views in doing so. Secondly, the move from 2 to 2** is not solely motivated by a narrow set of counterexamples. Rather, this specification is made to better capture common intuitions about well-being in a range of cases. I pointed out earlier that part of the appeal of desire theories is that they tie well-being to the subjective attitudes that we most closely identify with (section 2). Although desire theories do not take proximity to identity to non-instrumentally magnify the effects of a desire's fulfilment on well-being, these theories are made more plausible by entailing that the desires that we identify most with generally affect well-being more than other desires. Consequently, there is a principled reason for preferring this formulation of the view.

In addition to facing a larger explanatory burden, stability-adjusted views face two extra problems. Firstly, these views entail that, other things being equal, deliberately delaying the fulfilment of desires increases the amount of well-being that their fulfilment produces. For instance, if I can fulfil my desire to go for a run at any point during the day, then these views entail that deliberately delaying doing so until the evening is the more prudent choice. This can seem unintuitive. Yet, if the longevity principle is correct, then the conclusion follows.

While this conclusion may initially seem unintuitive, it is not unpalatable in all cases. Sometimes people have desires that specify a wide time preference for fulfilment. It strikes me that deliberately delaying the fulfilment of these desires may well increase the amount of well-being that their fulfilment produces. Take, for instance, the desire to climb Mt. Everest before one's death. Over time, what started as a flight of fancy may become a serious life goal that increasingly becomes part of one's sense of identity. Consequently, it does not strike me as intolerably unintuitive to accept that deliberately delaying the fulfilment of this desire does amplify the amount of well-being produced by its fulfilment. Therefore, we ought not to rule out the idea that well-being is improved by deliberately delaying the fulfilment of some desires. However, this is an unconvincing response for the bulk of cases. There are many desires that we aim to fulfil as quickly as possible. Take, for instance, the desire to drink water when thirsty. It is counterintuitive to claim that deliberately delaying the fulfilment of this desire increases the amount of well-being that its fulfilment produces.

In response to cases like this, we can claim that the non-instrumental benefit of deliberately delaying fulfilment is sometimes outweighed by the additional harms incurred by doing so. There are two ways in which this seems to be true. Firstly, deliberately delaying the fulfilment of a desire sometimes means frustrating other desires. Consider again the example of deliberately delaying the fulfilment of the desire to drink water when thirsty. According to stability-adjusted views, this non-instrumentally increases the amount of well-being that its fulfilment produces. Nevertheless, the delay also frustrates other desires. Most obviously, it frustrates the desire to be free from discomfort. On this view, the benefits of deliberately delaying the fulfilment of desires are sometimes outweighed by the additional harms incurred. Consequently, stability-adjusted views do not entail that it is always prudent to deliberately delay the fulfilment of desires.

There is a second way that these views can recognise that deliberately delaying the fulfilment of desires does not always improve well-being. This involves making a distinction between unfulfilled desires that are in a state of frustration and those that are in a state of anticipation.⁹ Sometimes our well-being is decreased by the absence of the things we desire. This happens when our desire aims at immediate fulfilment.

⁹Kris McDaniel and Ben Bradley have argued that there are additional states that desires can be in aside from fulfilment or frustration. They claim that desires can also be cancelled (2008: 274). This happens 'when a person's desire that P is conditional on Q, the desire that P is cancelled if and only if Q is false' (2008: 275). On their view, conditional desires with unsatisfied conditions are cancelled, rather than fulfilled. The idea that desire cancellation is an alternative state that desires can be in aside from fulfilment or frustration is a precedent for my claim that desires that specify a wide time preference are not frustrated until the window of time that they specify has elapsed.

Daniel Pallies has recently argued that desires are divisible into attractions and aversions. He argues that when we have an aversion to something, avoiding that thing does not non-instrumentally improve well-being. Conversely, when we have an attraction to something, failing to get that thing does not diminish well-being (Pallies 2022). This distinction allows him to explain how some desires operate purely positively, and others purely negatively, on well-being. If this is right, then there are additional resources available to

Take, for instance, a child's desire to open presents before Christmas Day. This desire is in a state of frustration for the length of time that it is held prior to fulfilment. While prolonging the length of time that this desire is held increases the amount of well-being produced by its fulfilment, doing so also decreases well-being by prolonging its frustration prior to that fulfilment. Therefore, delaying the fulfilment of desires that aim at immediate fulfilment does not improve aggregate well-being.

Conversely, some desires are in a state of anticipation before fulfilment. Take, for instance, the desire to go for a walk at any time before sunset. This desire specifies a wide time preference. Consequently, it is not frustrated until the end of the timeframe that it specifies. In cases like this, stability-adjusted views entail that deliberately delaying the fulfilment of the desire increases the well-being produced by its fulfilment without incurring countervailing harms. If this is correct, then stability-adjusted views can claim that deliberately delaying the fulfilment of some desires does improve well-being, while deliberately delaying the fulfilment of others does not. The way we distinguish between which desires have which effects depends upon whether the desire specifies a wide time preference or aims at immediate fulfilment. Consequently, stability-adjusted views have two ways of explaining why deliberately delaying the fulfilment of desires does not always improve well-being, and they can provide one argument in favour of the conclusion that sometimes deliberately delaying fulfilment does magnify the amount of well-being that a desire fulfilment produces.

There is a second problem worth considering for stability-adjusted views. This concerns the intuition that the strength of desire at the time of its fulfilment or frustration matters at least slightly more to well-being than the strength of desire at other times. Adopting the averaging principle means that the theory struggles to capture this intuition. To illustrate the intuitive force of this objection, consider the following two desires:

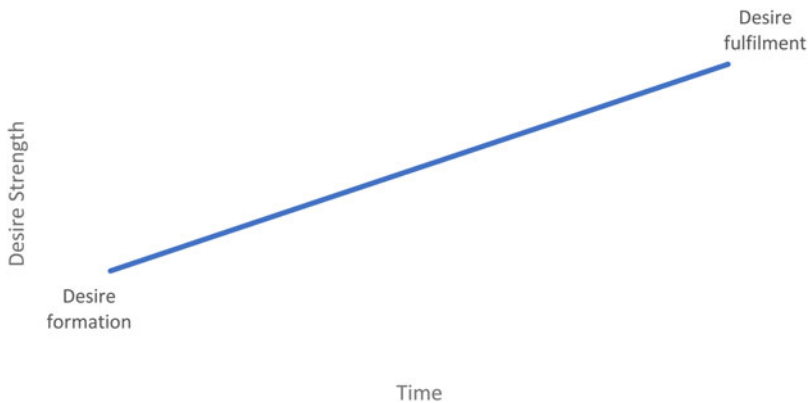


Figure 1. Ascending desire.

Ascending desire (see Figure 1) starts as a relatively weak desire. Over time it gradually becomes stronger. At the time of fulfilment, the desire is three times stronger than when it was formed.

desire theories for explaining why delaying the fulfilment of desires with a wide time preference does not diminish well-being.

Consider a second desire:

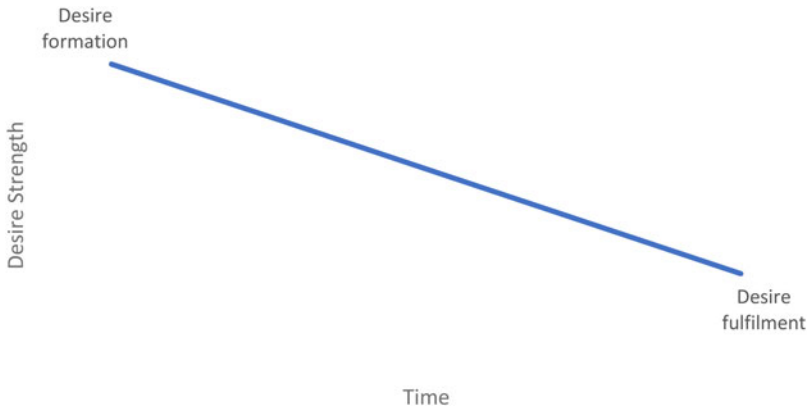


Figure 2. Descending desire.

Descending desire (see Figure 2) has the inverse shape to ascending desire. It starts strong. Over time it gradually becomes weaker. At the time of fulfilment, the desire is three times weaker than when it was formed.

Stability-adjusted views imply that these desires have identical effects on well-being. This is because both desires are held for the same length of time and have the same average strength. Nevertheless, there is some intuitive force in the idea that the fulfilment of ascending desire increases well-being at least slightly more than that of descending desire. Stability-adjusted views ought to have something to say about this intuition.

It is worth pointing out that not everyone shares this intuition. Those who find desire theories attractive may be especially prone to rejecting it. This is because these theories claim that a subject's well-being is not comprised solely of their mental states. Instead, they take well-being to be determined by a relationship between desire and the world (Bradley 2007: 47; Kagan 1994: 312).¹⁰ According to desire theories, a subject's well-being is improved when a desire corresponds to a state of the world in which it is fulfilled.¹¹ Conversely, their well-being is decreased when a desire corresponds to a state of the world in which it is frustrated. Notably, desires can sometimes be fulfilled without producing feelings of satisfaction and frustrated without producing feelings of frustration. Consider the case of an exile who is cut off from information about the lives of their children (Parfit 1984: 495). Assume that the exile retains a strong desire that their children live good lives, while lacking any information about them. Desire theories entail that whether this desire is fulfilled or frustrated affects the exile's well-being

¹⁰A possible counterexample is Chris Heathwood's subjective desire satisfactionism. This view claims that a subject's well-being is improved by the subjective perception that their desire has been fulfilled, and is decreased by the subjective perception that their desire has been frustrated. He takes this theory to be extensionally equivalent to attitudinal hedonism (Heathwood 2006: 559).

¹¹An alternative formulation claims that it is the states of affairs that we desire that are non-instrumentally valuable, rather than relational states of desire and world (Dorsey 2013: 152–53). The position that we take on this issue does not affect the wider arguments in this paper. Consequently, readers are free to replace my definition with this alternative if they prefer.

despite its failure to produce any beliefs or feelings within them. These theories also capture the intuition that well-being is improved less by misperceptions and deceptions than by phenomenologically identical veridical perceptions. Robert Nozick invites us to imagine a machine that simulates experiences in a way that is perceptually indistinguishable from reality. The machine is programmed to be optimal at generating positive mental states through its simulations (Nozick 1974: 42–3). Desire theories can claim that, despite our beliefs to the contrary, many of our desires are frustrated in the machine. This is because we are often not actually getting what we want when we mistake simulations for reality.¹² Consequently, these theories entail that we are better off if our feelings of satisfaction stem from fulfilled desires rather than mere simulacra.

The intuition that ascending desire matters more to well-being than descending desire appears to be based on the observation that fulfilling ascending desire often feels better than fulfilling descending desire. However, as we have seen, desire theories are already committed to rejecting the idea that feelings determine well-being. Indeed, this is one of the attractions of these theories. Consequently, those who find desire theories attractive are unlikely to be greatly troubled that these theories fail to capture the intuition that ascending desire matters more to well-being than descending desire. Nevertheless, the fact that fulfilling ascending desire often feels better may reveal one way in which it is more beneficial to us. It seems likely that we are the sort of creatures that have standing desires for feelings or attitudes such as pleasure or satisfaction. If this is right, then fulfilling ascending desire often improves well-being more than fulfilling descending desire because it better fulfils these standing desires. Consequently, stability-adjusted views can capture the intuition that ascending desire often improves well-being more than descending desire.¹³

We have seen that stability-adjusted views can navigate the problem of unstable desires and avoid the counterintuitive valuations made by simple concurrentism. They do so at the expense of acquiring the additional explanatory burden of how to balance desire strength with desire longevity. Moreover, these views are subject to two additional objections. The first one states that they counterintuitively entail that deliberately delaying the fulfilment of desires improves well-being. The second one states that these theories cannot capture the intuition that ascending desire matters more to well-being than descending desire. I have shown that stability-adjusted views can capture key intuitions in both cases.

¹²While this is frequently cited as a virtue of desire theories, it may be that these views do not fare much better than hedonism when it comes to the experience machine (Lowe and Stenberg 2017). This is because, while some desires are frustrated in the machine, it is possible that overall more desires are fulfilled than they would be in the real world. This outweighs the frustrations incurred in the machine. If this is right, then these theories may be forced to accept that it is prudent to plug into the experience machine.

¹³This discussion resembles a debate about whether the shape of a life non-instrumentally affects a subject's lifetime well-being. David Velleman argues that lives with an upward well-being trajectory are, other things being equal, non-instrumentally better for the subject living them than lives with a downward well-being trajectory (1991). Conversely, Fred Feldman argues that the shape of a life does not non-instrumentally affect the subject's lifetime well-being (2004: ch. 6). He points out that lives with an upward trajectory may be instrumentally better because many people prefer such lives. There are some parallels between Feldman's position on this debate and my analysis of ascending and descending desire. Nevertheless, intuitions about the importance of life trajectory do not necessarily translate into similar intuitions about whether ascending desire matters more than descending desire. The discussions may be parallel, but the issues are distinct.

5. Value-fulfilment theories of well-being

All good theories risk being supplanted by better ones. While stability-adjusted desire theories avoid the counterintuitive valuations made by simple concurrentism, their complexity may make them less attractive than rival theories. Some of the concerns that motivate the adoption of stability-adjusted desire theories are appealed to in support of value-fulfilment theories of well-being (henceforth, value-fulfilment theories).

The word ‘values’ can be suggestive of complex normative commitments. However, proponents of value-fulfilment theories often have in mind a wider conception of values than this. They consider values to be a subset of our pro-attitudes. That subset is restricted to pro-attitudes that conform to a specified set of requirements. Normally those requirements are largely concerned with how the pro-attitude relates to other aspects of our mental lives (Raibley 2010: 608). This allows us to conceive of values as a relatively broad set of pro-attitudes. Moreover, they need not be as cognitively sophisticated as the word ‘value’ sometimes implies (Yelle 2016: 1416). Consequently, these views should not be immediately dismissed for excluding too much of what we intuitively think non-instrumentally affects well-being.

A compelling version of the value-fulfilment theory is put forward by Jason Raibley. He argues that a subject’s well-being is primarily determined by the realisation of their values, and the cultivation of dispositions towards their realisation (2010: 596). He defines valuing as the possession of a pro-attitude towards a state of affairs that the subject stably identifies with (2010: 606–7). The extent to which the realisation of a value non-instrumentally improves well-being is dependent upon the pro-attitude’s intensity and the extent to which the subject identifies with it (Raibley 2010: 608). We may want to add to this view the requirement that values must be autonomously acquired in order for their realisation to improve well-being (Yelle 2014).

Raibley thinks that the pro-attitudes that we identify with normally have at least three components. Firstly, they have a phenomenology that is distinct from that of pro-attitudes alienated from our identity. Secondly, we are disposed towards approving of the pro-attitude and we take it as representative for who we want to be. Thirdly, we perceive the pro-attitude as reason-giving in the sense of justifying actions. It is the realisation of pro-attitudes that meet these criteria that primarily improves well-being. Raibley supplements his view by also noting that some affective states also non-instrumentally affect well-being (2010: 609).

Part of the reason that value-fulfilment theories are attractive is because our desires are not always reflective of our deepest concerns (Raibley 2010: 599; Yelle 2014: 372). Some desires are alienated from our identity despite resonating with other aspects of our psychology. Fleeting desires seem to be more frequently alienated from our identity. According to value-fulfilment theories, the fact that we are less likely to stably identify with these desires means that their fulfilment often does not benefit us (Raibley 2010: 614; Yelle 2014: 374). Consequently, these views are well-placed to avoid some of the counterintuitive valuations made by simple concurrentism. However, the subset of desires that are alienated from our identity is not limited to fleeting desires. Some people have the intuition that the fulfilment of non-fleeting desires that are alienated from our identity does not benefit us. If we share this intuition, then value-fulfilment theories have an advantage that stability-adjusted desire theories lack. They account for unstable desires *and* address concerns about alienation.

If value-fulfilment theories can better capture intuitions in these cases, then their appeal may eclipse that of stability-adjusted desire theories. However, there are good reasons to think that value-fulfilment theories are a less attractive family of views

about well-being. I focus on two arguments for this claim. The first highlights the larger explanatory burden that these theories incur. The second undermines the reasons for finding these theories attractive in the first place.

Firstly, value-fulfilment theories are more conceptually convoluted than desire theories. Not only must proponents of these theories construct a compelling theory of the valuing attitude, but they must also postulate that multiple things non-instrumentally affect well-being. Otherwise, the theory entails that beings incapable of valuing are also incapable of having well-being (Lin 2017: 357–65). This would exclude babies, many non-human animals, and some adult humans from being subjects of well-being.¹⁴ This would clearly be an unacceptable result. It can be avoided by postulating additional things that non-instrumentally affect well-being. However, adopting a multi-faceted theory of well-being means that value-fulfilment theories sacrifice significant conceptual parsimoniousness. Consequently, they incur a larger explanatory burden than the comparatively conceptually parsimonious stability-adjusted desire theories.

Secondly, the reasons for favouring value-fulfilment theories over desire theories are not particularly strong. If one is drawn to a value-fulfilment theory partly because of the counterintuitive valuations that many desire theories make about the importance of fleeting desires relative to long-standing desires, then stability-adjusted desire theories can avoid these valuations. Value-fulfilment theories and stability-adjusted desire theories differ on this issue because the former claims that many fleeting desires have no non-instrumental effects on our well-being, whereas the latter claims that these effects are simply less pronounced than those of long-standing desires. Even if one finds these cases more intuitively explained by value-fulfilment theories, it strikes me as unlikely that this benefit is worth the additional explanatory burden that these theories incur.

The main advantage that value-fulfilment theories have over desire theories is that the former can exclude desires that we do not identify with from affecting our well-being. Nevertheless, it is unclear to me whether we ought to exclude such desires. The strongest reason for doing so seems to be cases of addictions and compulsions. However, desire theories can explain cases of addiction as all-things-considered harms (Heathwood 2005: 493–94). These theories can claim that, while fulfilling desires born of addiction does indeed non-instrumentally benefit us, this benefit is outweighed by the countervailing harms incurred from doing so. These harms arise from the damage to health, finances, time, and relationships that addictions frequently result in. This frustrates more and stronger desires than the addiction typically fulfils. Consequently, we do not need to exclude addictive desires from non-instrumentally affecting well-being in order to explain why they are nevertheless bad for us.

It is possible to construct hypothetical examples whereby the addiction does not lead to these countervailing harms, does not produce any positive mental states in response to being fulfilled, but nevertheless does motivate us to acquire the addictive experience or substance (Parfit 1984: 497). I confess that it is counterintuitive to claim that fulfilling this type of addictive desire benefits us. Nevertheless, this type of example is not fatal to desire theories. We should be more sceptical of intuitions generated by extreme thought experiments than we are of real-world cases. Desire theories get the intuitive answer in

¹⁴An alternative approach claims that value-fulfilment theories only account for the well-being of subjects with the capacity to value (Yelle 2016). However, this approach inherits the unenviable task of explaining which standards apply to subjects that lack this capacity. Moreover, the view must accept that sometimes the standards of a subject's well-being change when they acquire or lose the capacity for valuing. This means that a subject's well-being could radically change for better or worse when they acquire or lose this capacity (Lin 2017: 359–60).

real-world cases of addiction. Moreover, such desires can still be all-things-considered harms providing that the subject experiencing them has a second-order desire to be free from them. It strikes me that most people in this situation would develop such desires. After all, the way that the thought experiment is constructed suggests that the subject gets nothing out of fulfilling the desire. It would be unusual for people to want to retain such desires. If the agent truly did not wish to be free from this desire, then it seems that the problem may be applicable to value-fulfilment theories as well. The lack of a second-order desire to be free from the addictive desire suggests a level of identification with it. For these reasons, we ought not to be too troubled by alleged counterexamples based on addiction.

Turning to compulsions, desire theories appear to suggest that fulfilling compulsions benefits us. This seems counterintuitive regarding compulsions that run counter to our values and interests. Compulsions can also be profoundly strange phenomena. Take, for instance, the call of the void. This is the compulsion to jump from a great height when experiencing vertigo (Schroeder 2004: 144). It is counterintuitive to claim that acting upon such a compulsion benefits us. Nevertheless, I do not think that the existence of compulsions fatally undermines desire theories. The all-things-considered harm explanation used to account for real-world addictions can also be applied to compulsions. Moreover, it may be that we mistakenly identify some compulsions as desires rather than alternative mental states. For instance, it may be that some compulsions are emotional responses to beliefs, rather than desires themselves. Consider again the vertiginous compulsion to jump from a great height. In this case, the phenomenology of the compulsion is similar to a desire, but the fact that it typically does not motivate us may suggest that we are not dealing with an actual case of desire. Of course, this seems less likely of compulsions that do motivate, but many things that are experienced as compulsions may, in fact, not be desires. Consequently, the existence of compulsions does not strike me as a sufficiently strong reason to accept a value-fulfilment theory.

Value-fulfilment theories can account for unstable desires and address concerns about alienation. Nevertheless, we have seen that there are good reasons to find these views less attractive than stability-adjusted versions of the desire theory. This is because these views inherit a larger explanatory burden and are not well-motivated. For these reasons, I have argued that the appeal of value-fulfilment theories does not eclipse that of desire theories.

Conclusion

Desire theories are popular views about well-being. However, the problem of unstable desires threatens to undermine their attractiveness. In this paper, I have sought to show that, far from undermining these theories, this problem can illuminate a version of the desire theory that better captures our intuitions about well-being. This view claims that the extent to which a desire fulfilment or frustration affects well-being is determined by the desire's average strength and magnified by the length of time that it is held. Clearly, there is a lot more to be said about these theories. Nevertheless, insofar as one finds desire theories compelling, stability-adjusted desire theories mark an improvement on how these theories are usually constructed.

Acknowledgements. I am grateful to Alex Gregory, Brian McElwee, Maria Paradeisopoulou, Kurt Sylvan, three anonymous reviewers, and the editor of *Utilitas*, Ben Eggleston, for comments on earlier versions of this paper. This material has also benefited from audience feedback at a postgraduate seminar at the University of Southampton in 2021. It was made possible by generous funding from the Royal Institute of Philosophy's Jacobsen scholarship.

Competing interests. None.

References

- Bradley, Ben.** 2007. A Paradox for Some Theories of Welfare. *Philosophical Studies* 133(1): 45–53.
- Bruckner, Donald W.** 2013. Present Desire Satisfaction and Past Well-Being. *Australasian Journal of Philosophy* 91(1): 15–29.
- Bykvist, Krister.** 2003. The Moral Relevance of Past Preferences. In *Time and Ethics: Essays at the Intersection*, ed. Heather Dyke. Dordrecht: Kluwer Academic Publishers, pp. 115–36.
- Dorsey, Dale.** 2013. Desire-Satisfaction and Welfare as Temporal. *Ethical Theory and Moral Practice* 16(1): 151–71.
- Feldman, Fred.** 2004. *Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism*. Oxford: Clarendon Press.
- Gregory, Alex.** 2021. *Desire as Belief: A Study of Desire, Motivation, and Rationality*. Oxford: Oxford University Press.
- Griffin, James.** 1986. *Well-Being: Its Meaning, Measurement and Moral Importance*. Oxford: Clarendon Press.
- Hare, R. M.** 1981. *Moral Thinking: Its Levels, Method and Point*. Oxford: Clarendon Press.
- Heathwood, Chris.** 2005. The Problem of Defective Desires. *Australasian Journal of Philosophy* 83(4): 487–504.
- Heathwood, Chris.** 2006. Desire Satisfactionism and Hedonism. *Philosophical Studies* 128(3): 539–63.
- Heathwood, Chris.** 2015. Desire-Fulfillment Theory. In *The Routledge Handbook of Philosophy of Well-Being*, ed. Guy Fletcher. London and New York: Routledge, pp. 135–47.
- Heathwood, Chris.** 2019. Which Desires are Relevant to Well-Being? *Noûs* 53(3): 664–88.
- Hubin, Donald C.** 2003. Desires, Whims and Values. *The Journal of Ethics* 7(3): 315–35.
- Kagan, Shelly.** 1994. Me and My Life. *Proceedings of the Aristotelian Society, New Series* 94: 309–24.
- Lauinger, William.** 2011. Dead Sea Apples and Desire-Fulfillment Welfare Theories. *Utilitas* 23(3): 324–43.
- Lewis, David.** 2000. Dispositional Theories of Value. In *Papers in Ethics and Social Philosophy*. Cambridge: Cambridge University Press, pp. 68–94.
- Lin, Eden.** 2017. Against Welfare Subjectivism. *Noûs* 51(2): 354–77.
- Lin, Eden.** 2018. Welfare Invariabilism. *Ethics* 128(2): 320–45.
- Low, Dan and Stenberg, Joseph.** 2017. The Experience Machine Objection to Desire Satisfactionism. *Journal of the American Philosophical Association* 3(2): 247–63.
- Mariqueo-Russell, Atus.** 2023. Desire and Motivation in Desire Theories of Well-Being. *Philosophical Studies* 180(7): 1975–94.
- McDaniel, Kris and Bradley, Ben.** 2008. Desires. *Mind, New Series* 117(466): 267–302.
- Murphy, Mark C.** 1999. The Simple Desire-Fulfillment Theory. *Noûs* 33(2): 247–72.
- Nozick, Robert.** 1974. *Anarchy, State, and Utopia*. Malden, MA, Oxford and Carlton, OR: Blackwell Publishing.
- Pallies, Daniel.** 2022. Attraction, Aversion, and Asymmetrical Desires. *Ethics* 132(3): 598–620.
- Parfit, Derek.** 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Pineda-Oliva, David.** 2021. Defending the Motivational Theory of Desire. *Theoria* 36(2): 243–60.
- Pitcher, George.** 1984. The Misfortunes of the Dead. *American Philosophy Quarterly* 21(2): 183–88.
- Raibley, Jason.** 2010. Well-Being and the Priority of Values. *Social Theory and Practice* 36(4): 593–620.
- Railton, Peter.** 1986. Facts and Values. *Philosophical Topics* 14(2): 5–31.
- Rosati, Connie.** 1996. Internalism and the Good for a Person. *Ethics* 106(2): 297–326.
- Sarch, Andrew.** 2013. Desire Satisfaction and Time. *Utilitas* 25(2): 221–45.
- Schroeder, Mark.** 2007. *Slaves of the Passions*. Oxford: Oxford University Press.
- Schroeder, Timothy.** 2004. *Three Faces of Desire*. New York: Oxford University Press.
- Sidgwick, Henry.** 1981 [1907]. *The Methods of Ethics*. 7th ed. Indianapolis, IN: Hackett Publishing Company.
- Velleman, David J.** 1991. Well-Being and Time. *Pacific Philosophical Quarterly* 72(1): 48–77.
- Vorobej, Mark.** 1998. Past Desires. *Philosophical Studies* 90(3): 305–18.
- Yelle, Benjamin.** 2014. Alienation, Deprivation, and the Well-Being of Persons. *Utilitas* 26(4): 369–84.
- Yelle, Benjamin.** 2016. In Defense of Sophisticated Theories of Welfare. *Philosophia* 44(4): 1409–18.