Against Welfare Subjectivism

Eden Lin
Rutgers University, Newark

Subjectivism about welfare is the view that something is basically good (bad) for you if and only if, and to the extent that, you have the right kind of favorable (disfavorable) attitude toward it under the right conditions. I make a presumptive case for the falsity of subjectivism by arguing against nearly every extant version of the view. My arguments share a common theme: theories of welfare should be tested for what they imply about newborn infants. Even if a theory is intended to apply only to adults, the fact that it is false of newborns may give us sufficient reason to reject it.

Some things are *basically* good for you—good for you, but not solely in virtue of being suitably related to other things that are good for you. *Subjectivists* about welfare claim that something is basically good for you if and only if (and to the extent that), in a certain familiar sense, it *suits* or *fits* you. Some things suit certain people better than others: the symphony better suits aficionados of classical music than it does people who only like jazz. Some things do not suit any normal human being: drinking a can of paint, or spending a year counting blades of grass. Something suits you, in this familiar sense, to the extent that you have (or would, in the right conditions, have) a *favorable attitude* toward it. Thus, subjectivism says that something is basically good for you just if, and to the extent that, you have a certain favorable attitude, \(A\), toward it under the right conditions, \(C\).

Determinate versions of subjectivism specify \(A\) and \(C\). For instance, on one version, something is basically good for you if and only if and to the extent that you would *desire* it if you were *fully informed and perfectly rational*.\(^1\)

Extant arguments against subjectivism focus on particular formulations of it, leaving it open that other versions are viable.\(^2\) Indeed, all existing discussions assume an interpretation of attitude \(A\) that

---

\(^1\) Subjectivism is an explanatory thesis: it says that \(x\)'s degree of basic goodness for you is *due to* your having \(A\) toward \(x\) to a certain extent in \(C\). For the sake of brevity, I will use ‘\(P\ iff\ Q\)’ to stand for ‘\(P\ iff, because, and to the extent that\) \(Q\)’. Note, too, that the subjectivist sense of suitability or fit is distinct from the perfectionist sense invoked by Kraut (2007). For discussion of suitability in the subjectivist sense, see Rosati (1995b), p. 50 and (1996) pp. 298-99.

ignores a type of subjectivism that has recently been defended. I aim to argue against nearly every subjectivist view with contemporary defenders. First, I give a taxonomy of subjectivist views by distinguishing three specifications of $A$ and two specifications of $C$. Then, I argue against all six of the resulting views. Since practically every existing subjectivist theory is one of these views or a combination thereof, this amounts to a strong presumptive case against subjectivism.

My arguments share a common theme: views about welfare should be tested for what they imply about newborn infants. A view may appear more credible than it is because its implausible implications about neonatal welfare have gone unnoticed. Moreover, even if a view is intended to apply only to adults, we may have sufficient reason to reject it because it implies an implausible divergence between the welfare of adults and that of newborns.

1. Preliminaries

Before I state the subjectivist views that I will be discussing, a few clarifications are in order.

Although I will usually elide this, subjectivism is really the conjunction of the aforementioned claim about basic goodness and the corresponding claim about basic badness: something is basically bad for you if and only if you have a certain unfavorable attitude, $A^*$, toward it under conditions $C$—where $A^*$ is the opposite of $A$ (e.g., aversion, if $A$ is desire). Moreover, when I speak about whether something is basically good or bad for you, I am talking about whether some particular thing (e.g., a particular pleasure) has that status: I am not talking about whether some kind (e.g., pleasure) is a basically good or bad kind—or as I like to call it, a basic good. Your level of welfare is fixed by how basically good or bad for you all of the particular things are. Subjectivism says that these facts are wholly determined by the facts about the favorable and unfavorable attitudes that you have toward those things under the right conditions.

---

3 Namely, the one defended in Dorsey (2010) and (2012).
4 Kraut (2007) endorses this principle. Rosati (2009) notes that some theories of welfare appear not to apply to infants, but she doesn’t say that we should therefore reject them.
5 For discussion of basic goods and their relation to basically good particulars, see Lin (2015).
6 Although subjectivism says that a particular thing is basically good for you iff you have $A$ toward it in $C$, a particular can satisfy this condition by belonging to a type toward which you have $A$ in $C$. You might have a desire for admiration that involves desiring, of any arbitrary person, that he admires you.
I assume that according to subjectivism, the basically good things are the objects of \( A \) in \( C \) (e.g., the states of affairs that you would desire if fully informed). But some subjectivists prefer to say that the basically good things are combinations of those objects and your having \( A \) toward them in \( C \) (e.g., states of the form \( p \& \text{you would desire } p \) if fully informed). Although I assume an ‘object’ formulation of subjectivism, my arguments could be translated into the language of ‘combo’ subjectivism.

Whether hedonism is a subjectivist view depends on the relation between pleasure and favorable attitudes. Suppose that a pleasure is just a mental episode in which, among other things, you have a certain favorable attitude (e.g., desire) toward some state of affairs. Then since hedonism says that all and only pleasures are basically good, it is a subjectivist theory of the ‘combo’ variety. On the other hand, suppose that you can feel a pleasure without having a favorable attitude toward anything. Then since hedonism says that all pleasures are basically good, it isn’t a subjectivist theory. Since I remain neutral on the metaphysics of pleasure, I remain neutral on how to classify hedonism. Because I think hedonism faces serious problems whether or not it is a subjectivist view—problems distinct from the ones that I will be raising against subjectivism—I will not discuss hedonism in this paper.

Finally, I understand subjectivism to be a view in normative ethics, not metaethics. It concerns which things have the property of basic goodness for you (and why), not what this property is. Thus, it is neutral on whether this property is irreducibly normative, and on whether it is natural. Two non-naturalists could disagree about whether something has the non-natural property of basic goodness for you if and only if you have \( A \) toward the thing in \( C \).

2. A Taxonomy of Subjectivist Views

A theory of welfare shouldn’t merely tell us what is basically good for you at the actual world: it should also tell us what is basically good for you at other possible worlds. If you aren’t feeling happy

---

7 See, e.g., Bradley (2014) and Heathwood (2014).
8 Again, see Bradley (2014) and Heathwood (2014).
10 This view is held by Bramble (2013) and perhaps also by Smuts (2011).
11 See Lin (2014) and (forthcoming a) for discussions of hedonism. The argument against Same World Desire Subjectivism that I give below would also count against hedonism.
right now, a theory of well-being should tell us whether it would have been basically good for you to feel happy now—i.e., whether at the closest world where you feel happy now, this is basically good for you. For any world, a theory of well-being should tell us what (if anything) is basically good for you there. (Plausibly, if you do not exist at a world, then nothing is basically good for you there.)

Subjectivism says that something $x$ is basically good for you at $W$ if and only if you have $A$ toward it in $C$. There are two salient candidates for $C$: it could be the same world at which the basic goodness for you of $x$ is in question (viz., $W$), or it could be the ideal world—the closest world at which you are idealized in the sense of being fully informed and fully instrumentally rational. Thus, there are two salient kinds of subjectivist view that differ on what the right conditions are:

**Same World Subjectivism**

$x$ is basically good for you at $W$ if and only if at $W$, you have favorable attitude $A$ toward $x$.

**Ideal World Subjectivism**

$x$ is basically good for you at $W$ if and only if at the ideal world, $I$, you have favorable attitude $A$ toward $x$—where $I$ is the closest world where you have full information (including firsthand acquaintance with all possible experiences) and full instrumental rationality.

Some subjectivists claim that what matters is what attitudes you have at $I$ toward your wanting, pursuing, or having $x$ at $W$.

I intend Ideal World Subjectivism to be read in such a way that such attitudes are, broadly speaking, attitudes toward $x$. Thus, I classify these theorists as Ideal World Subjectivists.

There are three obvious candidates for attitude $A$: a belief or judgment that the thing is basically good for you, a desire, or a valuing. This yields a threefold distinction that is orthogonal to the one between Same World and Ideal World subjectivism:

---

12 In this paper, I assume the standard possible worlds semantics for counterfactual conditionals. For simplicity, I also assume that for any counterfactual, there is a closest world where its antecedent is true.


Judgment Subjectivism: $x$ is basically good for you at $W$ if and only if under the right conditions, $C$, you believe that $x$ is basically good for you at $W$.\(^{15}\)

Desire Subjectivism: $x$ is basically good for you at $W$ if and only if under the right conditions, $C$, you desire $x$.\(^{16,17}\)

Value Subjectivism: $x$ is basically good for you at $W$ if and only if under the right conditions, $C$, you value $x$.

I assume that valuing something involves having a favorable conative attitude toward it that is more psychologically complex than desire. Perhaps you value $x$ just if you desire to desire it.\(^{18}\) Or perhaps you value $x$ just in case you have a pro-attitude toward it with which you identify over a long period of time—where you identify with an attitude just if you don’t experience it as ‘alien’ to you, it doesn’t conflict with your other attitudes about its object, you are disposed to approve of it, and you are disposed to treat it as giving you reasons to act in certain ways.\(^{19}\) Without endorsing any particular account of valuing, I assume that some theory of this sort is correct and that desiring something is not sufficient for valuing it. These assumptions are widely-held.\(^{20,21}\)

Desire Subjectivism is the most common form of subjectivism. Judgment and Value Subjectivism are relative newcomers. Dale Dorsey defends Ideal World Judgment Subjectivism\(^{22}\), and Benjamin

---

\(^{15}\) Why isn’t this view viciously circular? Because it is merely a view about which things are basically good and why, not an analysis of basic goodness. In this respect, it is similar to Sharon Street’s constructivist view that $X$ is a reason for agent $A$ to $Y$ if and only if (and because) “the judgment that $X$ is a reason to $Y$ (for $A$) withstands scrutiny from the standpoint of $A$’s other judgments about reasons” (2008, p. 223). Just as we can take the notion of a reason as primitive and explain which things are reasons in terms of judgments about reasons, we can take the notion of prudential value as primitive and explain which things have such value in terms of judgments about prudential value. See Street (2008), pp. 239-42 and Dorsey (2012), pp. 438-41. If these considerations do not exonerate Judgment Subjectivism from the charge of vicious circularity, then so much the worse for subjectivism.

\(^{16}\) Most desire subjectivists would say that $A$ is desire of a certain kind—e.g., intrinsic desire, or perhaps desire in the “genuine attraction” sense (see Heathwood (unpublished)). I remain neutral on what the right kind of desire is. When I speak of desire, you should read me as speaking of desire of the right kind, whatever that turns out to be.

\(^{17}\) On Railton’s formulation of Ideal World Desire Subjectivism, the object of the desire is not $x$, but your desiring or pursuing $x$. It would be tedious to complicate the formulation of Ideal World Desire Subjectivism to reflect this fact.

\(^{18}\) Lewis (1989).

\(^{19}\) Raibley (2010), pp. 606-08.


\(^{21}\) Dorsey (2012) argues that to value something is (roughly) to believe that it is good for you, so he would equate Judgment and Value Subjectivism. For simplicity, I am stipulating a use of ‘value’ and its cognates that rules this out.

\(^{22}\) Dorsey (2012).
Yelle defends a ‘combo’ version of Same World Value Subjectivism.\textsuperscript{23} Though Judgment and Value Subjectivism disagree about what $A$ is, they agree that it is an attitude that requires a fair amount of psychological sophistication. We can therefore call them sophisticated subjectivist views.

The foregoing specifications of $A$ and $C$ generate six views. I will first give an argument against the two sophisticated Same World views (viz., Same World Judgment Subjectivism and Same World Value Subjectivism). Then, I will give a different argument against Same World Desire Subjectivism. Finally, I will give an objection to all Ideal World subjectivist views, however they specify $A$.

3. Sophisticated Same World Subjectivist Views

In this section, I give an argument against the two sophisticated Same World subjectivist views. I will first run the argument against

\begin{align*}
\text{Same World Subjectivism} & : x \text{ is basically good for you at } W \text{ if and only if at } W, \text{ you believe that } x \text{ is basically good for you.} \\
\text{Judgment} & : x \text{ is basically good for you.}
\end{align*}

Then, I will explain how the same argument could be run against Same World Value Subjectivism.

Let us call something a welfare belief just if it is a belief about whether something is good or bad for you. Same World Judgment Subjectivism entails that something is basically good for you at a world only if it is good for you according to the welfare beliefs you have there.\textsuperscript{24} It therefore implies that if, at $W$, you do not have any welfare beliefs, then nothing is basically good for you at $W$.\textsuperscript{25} Clearly,

\begin{itemize}
\item \textsuperscript{23} Yelle (2014) claims that your welfare consists in your self-realization, which, in turn, consists in the realization of your values. So his view is the ‘combo’ subjectivist view that the basically good states of affairs are ones in which you realize something you value. He claims (p. 373) that his view is objectivist, but it seems to me that he has neglected the possibility of ‘combo’ subjectivist views.
\item \textsuperscript{24} Dorsey (2012, pp. 415-16) worries that your welfare beliefs at $W'$ might be incoherent in a way that implies that $x$ both is and isn’t basically good for you at $W'$. He therefore suggests that we look not at your welfare beliefs at $W'$, but at your welfare beliefs at the closest world to $W'$.\textsuperscript{25} Their coherence is purged in accordance with the principle of “minimal mutilation.” But since it is not possible for you to both believe and fail to believe that $x$ is basically good for you, I don’t share this worry. My argument would apply equally to Dorsey’s formulation of the view. (See the next note.)
\item \textsuperscript{25} Dorsey’s reformulation has the same implication, since if you lack welfare beliefs at $W'$, then the closest world to $W'$ where you don’t have incoherent welfare beliefs is $W'$. Attempting to render an empty set of welfare beliefs coherent in accordance with the principle of minimal mutilation can’t yield a non-empty set of such beliefs.
\end{itemize}
however, there are entities for whom certain things are basically good at worlds where they lack welfare beliefs. Newborn babies are like this. They are clearly capable of well-being. If you hit a newborn, it will be worse off than it otherwise would have been. If you feed it when it is hungry or you cuddle it, it will be better off. Indeed, it is clear that a newborn can have a positive level of welfare, and this implies that some things can be basically good for it. Newborns don’t have welfare beliefs, however. Although recent advances in developmental psychology have shown that older babies are surprisingly sophisticated, there is no evidence that newborn babies have welfare beliefs. They like some things and are averse to others, and perhaps they have beliefs. But they surely do not have beliefs to the effect that X is good for them. Moreover, even if we can’t be certain that this is actually true, we can be certain that there are many possible worlds at which this is true, and that many of these are worlds where some things are basically good for certain newborns. Thus, the following claim is true:

**Babies**

There are possible worlds at which something is basically good for a newborn even though the newborn does not have any welfare beliefs at that world. (It is overwhelmingly likely that the actual world is such a world.)

This implies that Same World Judgment Subjectivism is false—at least of newborns. For that view implies that there are no such worlds.

I assumed that, according to Same World Judgment Subjectivism, the time at which you must believe that $x$ is basically good for you is the time at which $x$ is basically good for you. What if we interpret it as allowing your future welfare beliefs to account for a thing’s present basic goodness? My argument would still go through, even though humans typically have welfare beliefs once they reach a certain age. After all, something can be basically good for a newborn who does not survive long enough to acquire any welfare beliefs. These newborns make Babies true even when its last clause is read without an implicit restriction to the time at which $x$ is basically good for $S$. And since nothing that is basically good for these tragically short-lived newborns has that status because they will believe it to be good for them in the future, the same is surely true of newborns who will survive long enough to have welfare beliefs. Since Same World Judgment Subjectivism is false of newborns who
die before they acquire any welfare beliefs, it is also false of newborns who will acquire such beliefs in the future.

Now, Same World Judgment Subjectivism might be read as a view about the welfare of cognitively normal human adults rather than as a thesis about all welfare subjects: for although it is false of subjects that lack the capacity to have welfare beliefs (e.g., newborns, severely mentally disabled adults), it might nonetheless be true of all subjects that possess that capacity. But even if the view is restricted in this way, we should reject it. For as I will now argue, if the view is true of normal adults even though it is false of newborns, then adult welfare diverges from neonatal welfare in a way that cannot plausibly be explained.

If Same World Judgment Subjectivism is true of adults, then whenever something is basically good for an adult, what makes this true is the fact that she believes it to be basically good for her. But as we have seen, this is not how things work with newborns. Consider an adult who didn’t have welfare beliefs when she was a newborn. Why would some fact about her welfare beliefs explain why anything is basically good for her now, even though no such fact explained why anything was basically good for her when she was a newborn? Why would a test that everything that was basically good for her previously failed to pass be one that everything that is basically good for her must now pass? Because Same World Judgment Subjectivism is false of newborns, the following principle creates a presumption that it is also false of adults:

**Difference**

If X isn’t the explanation of why everything that is basically good for subjects of kind A is basically good for them, then it isn’t the explanation of why everything that is basically good for subjects of kind B is basically good for them—unless there is some difference between As and Bs that explains why they would differ in this way with respect to X.

To defeat this presumption, one must identify a difference between normal adults and newborns that explains why Same World Judgment Subjectivism could be true of the former even though it is false of the latter. Although newborns and normal adults differ extensively, it is difficult to see how this could be done.
If Same World Judgment Subjectivism is true of normal adults even though it is false of newborns, that is surely because it becomes true of a human being when she acquires the capacity to have welfare beliefs. It might be argued that when someone acquires this capacity, she becomes an importantly different kind of creature: one capable of having a conception of her own good. Her nature has changed dramatically, and this explains why Same World Judgment Subjectivism is now true of her. But this is unpersuasive. Although the capacity to have a conception of one’s own good is a major difference between normal adults and newborns, it is hard to see how the fact that a normal adult has this capacity could plausibly explain why there is a restriction on the things that can be basically good for her that was not in place when she was a newborn. It is one thing to suppose that when someone acquires this capacity, more things are potentially basically good for her than previously were. This particular consequence of Same World Judgment Subjectivism seems explicable. For it is at least somewhat plausible that by believing something to be basically good for her, someone can confer basic prudential value on it. By contrast, I don’t see what could plausibly explain the alleged fact that when someone acquires the capacity to have welfare beliefs, things that were previously basically good for her might now be barred from having this status simply because she fails to believe them to have this status. Many things were previously basically good for her even though she didn’t believe them to be. Why would they suddenly cease to be basically good for her simply because she acquired the relevant capacity and her beliefs about them remained unchanged? The fact that she now has this capacity is insufficient to explain why there is a new necessary condition on basic goodness for her.

To see that this is so, consider a human being who acquires the capacity to have welfare beliefs at the normal stage in her development, but who fails, for a period of time after acquiring the capacity, to actually believe anything to be basically good for her. (This isn’t farfetched. Moore claimed not to have any welfare beliefs, since he claimed not to understand how something could be good for someone, as opposed to being good simpliciter and appropriately related to that person. Even a person who has the concept of basic prudential value might lack any welfare beliefs because exposure to conflicting theories has led her to suspend judgment on these matters.) Imagine that, prior to acquiring the relevant capacity, our subject is high in welfare. If acquiring this capacity makes Same World Judgment Subjectivism true of her, then from the time at which she acquires the

---

capacity until the later time at which she begins to believe things to be basically good for her, her welfare is at best zero: nothing is basically good for her during this interval, since she doesn’t believe anything to be basically good for her. This is incredible. If she hasn’t acquired a single welfare belief, and if she was previously high in welfare even though she lacked any such beliefs, then why would the mere fact that she is now capable of having such beliefs make it true that nothing is basically good for her? It is implausible, and it seems inexplicable, that a subject’s welfare could go from being high to being zero (or negative) simply because she has acquired a capacity that she previously lacked. After all, whatever conditions accounted for her previously high level of welfare remain in place: the only thing that has changed is that she has acquired the capacity.

To make this more vivid, imagine that hedonism is the correct theory of the welfare of human beings who cannot have welfare beliefs. (This supposition is consistent with the view that Same World Judgment Subjectivism is true of normal adults, all of whom can have such beliefs.) Our imagined subject is high in welfare until the point at which she acquires the capacity to have welfare beliefs. Because hedonism is true of her until then, she has a very favorable balance of pleasure over pain until that time. Now, imagine that her life remains hedonically unchanged even after the point at which she acquires the relevant capacity. If Same World Judgment Subjectivism becomes true of her at this point, then her welfare drops to zero (or lower) at that time—even though she continues to have exactly the same balance of pleasure over pain in virtue of which she was previously high in welfare. This is implausible. If those favorable hedonic conditions were formerly sufficient for a high level of welfare, they are surely still sufficient for a slightly positive level of welfare. For what could explain why they are suddenly of no benefit at all?

A similar argument could be made, no matter what theory we suppose is true of human beings who can’t have welfare beliefs, and no matter what capacity the acquisition of which is alleged to make Same World Judgment Subjectivism true of a welfare subject. It might initially seem that, since normal adults differ from newborns about as extensively as they do from dogs and other animals, this theory could be true of normal adults even though it is false of newborns. But this impression is mistaken. Because newborns eventually mature into adults, the claim that this view is true of normal

---

27 If you think that a subject for whom nothing is basically good or bad has no level of welfare (rather than a welfare level of zero), then you should read me as claiming that our subject’s level of welfare is either negative or undefined. I thank an anonymous referee for alerting me to this point.
adults implies that the welfare of normal adults diverges from that of newborns in a way that cannot plausibly be explained.

The Same World Judgment Subjectivist might respond that his view is true, not of all creatures who are capable of having welfare beliefs, but only of those who have such beliefs. But this would be to concede that his view is not even true of normal human adults. For even if every normal adult actually has welfare beliefs, for every such adult, there are worlds at which he never acquired any such beliefs after acquiring the capacity to have them. As I just argued, some of these worlds are ones where certain things are basically good for him when he is an adult. These worlds are counterexamples to Same World Judgment Subjectivism even when it is read as a view only about the welfare of normal adults. For whatever subjects this view is meant to apply to, the view purports to a state the explanation of anything’s being basically good for such a subject at any possible world.

The Same World Judgment Subjectivist might retreat even further to the claim that his view is true of normal adults only at those worlds where they believe at least one thing to be basically good for them. But this response will not do. For as this response concedes, a pleasure can be basically good for a normal adult, Ronnie, at a world \( W \) where Ronnie doesn’t believe anything to be good for him. And clearly, it could be that at a different world \( W^* \), that very pleasure is basically good for Ronnie and believed by him to be basically good for him. But surely, it could be that \( W \) is the closest world to \( W^* \) at which Ronnie doesn’t believe that that pleasure is basically good for him. If so, then at \( W^* \), it is true that the pleasure would be basically good for Ronnie even if he didn’t believe it to be basically good for him. Thus, at \( W^* \), it is false that the pleasure is basically good for Ronnie if and only if and because he believes it to be basically good for him. So Same World Judgment Subjectivism is not true of adults even at all the worlds where they believe at least one thing to be basically good for them.

To summarize: Same World Judgment Subjectivism is false of newborns because some things are basically good for them even though they lack welfare beliefs. Thus, we should conclude that it is also false of normal adults. For if it were true of normal adults, there would be an implausible and inexplicable divergence between adult and neonatal welfare, and someone could go from being high in welfare to having negative or zero welfare simply by reaching a certain stage of development (e.g., acquiring the capacity to have welfare beliefs).
This argument could be run against any Same World Subjectivist view according to which \( A \) is too psychologically sophisticated for newborn babies to have. It is not necessary that \( A \) be a cognitive attitude. Thus, the argument also rules out

**Same World Value** \( x \) is basically good for you at \( W \) if and only if at \( W \), you value \( x \).

**Subjectivism**

For whatever valuing turns out to involve, newborns lack the psychological sophistication required to value anything. For example, they don’t have higher order desires and are not disposed to treat any of their attitudes as giving them reasons for action. Certainly, this is true at some worlds where they are positive in welfare. Same World Value Subjectivism falsely implies that, at those worlds, nothing is basically good for them. Thus, it is false of newborns. It is difficult to see how one could explain why this view might nonetheless be true of normal adults: after all, it can’t be true of us but false of newborns unless someone can go from being high in welfare to being neutral or negative in welfare simply by acquiring the capacity to value things. We should therefore reject Same World Value Subjectivism even as a theory of the welfare of normal human adults.\(^{28}\)

4. **Same World Desire Subjectivism**

Now consider the desiderative version of Same World Subjectivism, which is just the simplest version of the desire satisfaction theory of welfare:

**Same World Desire** \( x \) is basically good (bad) for you at \( W \) if and only if, because, and to the extent that at \( W \), you desire (are averse to) \( x \).

Mark Murphy and Chris Heathwood can both be read as arguing that this is the best form of desire satisfactionism.\(^{29}\)

---

\(^{28}\) Some philosophers (e.g., Sumner (1995)) use ‘subjectivism’ to refer to something weaker than subjectivism in my sense—namely, the view that \( x \) is basically good for you only if (and at least partly because) you have \( A \) toward it in \( C \). My argument in this section also counts against sophisticated Same World subjectivism of this weaker variety.

\(^{29}\) Murphy (1999) and Heathwood (2005) independently claim that, according to the best form of desire satisfactionism, the satisfaction of your “actual” desires is what is basically good for you. If they are referring to a view on which what's basically good for you at any world is what you actually desire at that world, then they are endorsing Same World Desire Subjectivism. But if they have in mind a view on which what's basically good for you at any world is what you desire at
The kind of argument I just gave cannot be used against Same World Desire Subjectivism. Newborn babies have desires. They like certain things and dislike other things. They are motivated to obtain certain things and to avoid other things. These all count as desires in the thin sense standardly used in discussions of welfare. And it is not obvious that some things are basically good for newborns even at worlds where they lack desires. So we cannot reject Same World Desire Subjectivism on the grounds that newborns lack the psychological machinery that it deems necessary for well-being. However, this view runs afoul of a different, more particular fact about neonatal welfare.

Let us call a desire experiential just if it is a desire for an experience that is based solely on the felt quality of that experience. Surely, given how primitive they are psychologically, newborns only have experiential desires: they desire experiences that feel certain ways, they are averse to experiences that feel certain other ways, and those are the only kinds of desire they have. For how could a newborn desire not merely the cuddly sensations that he is experiencing, but (say) that they have a particular cause (e.g., his mother)? This would require an understanding of causation and of the world as being populated by causes—something that is, presumably, outside the ken of an infant who has just emerged from the womb.\(^{30}\)\(^{31}\) Now, consider two newborns who are duplicates as far as their beliefs and desires are concerned and who have experientially indistinguishable lives. The crucial difference between them is that whereas one of them (call him Adam) lives in the real world, the other one (call him Bill) is always plugged into an experience machine.\(^{32}\) Adam has the sort of life available to an infant born into a well-to-do family: he is raised by two loving parents, he begins to explore the world around him, and he learns to control and move his body in the normal ways. Bill has none of these things, but he is fed a simulation of them via the experience machine. (Just as he emerges from

---

The actual world, then they are endorsing a different view, Actual World Desire Subjectivism. That view is implausible: it implies that if you had completely different tastes and interests because you had a completely different upbringing, it would benefit you to obtain the things that you actually desire, rather than the things that you would desire. For more on the contrast between Same World Subjectivism and Actual World Subjectivism, see Lin (unpublished a).

\(^{30}\) As I will explain later, the assumption that newborns only have experiential desires is dispensable.

\(^{31}\) I am not aware of any research in developmental psychology that clearly suggests that newborn infants have desires that are not merely experiential. Meltzoff and Moore (1977) observed that newborns tend to match certain facial expressions that they are presented with, and some researchers (e.g., Gopnik (2009), ch. 9) have taken this to suggest that newborns understand the world as containing other conscious subjects and desire to imitate those subjects. But this interpretation is highly controversial, and a simpler explanation is available: see Jones (1996), (2006), and (2009).

\(^{32}\) See Nozick (1974), pp. 42-45. Nowadays, many philosophers are dismissive of arguments involving the experience machine. But their criticisms are directed at arguments in which the crucial premise concerns whether we would or should plug into the machine. The present argument isn’t undermined by their criticisms because it appeals to no such premise. I discuss these issues in Lin (forthcoming a).
the womb, robots kill his parents and plug him into the machine.) Is Adam’s total welfare during the neonatal period higher than Bill’s?

It is important to distinguish this question from that of whether, considering their entire lifetimes (which, we may suppose, span several decades), Adam and Bill are equal in total welfare. Many would say that Adam’s total lifetime well-being is higher. Same World Desire Subjectivism can accommodate this claim. After all, when Adam and Bill reach a certain stage of development, they most likely acquire non-experiential desires—desires concerning the real world. In so doing, they acquire a profile of desires that is better satisfied in Adam’s life than in Bill’s. For example, they both come to desire to have friends, and while this desire is satisfied in Adam’s case, it is frustrated in Bill’s. Thus, according to Same World Desire Subjectivism, Adam is higher in lifetime well-being. This is why the traditional case of the experience machine, focusing as it does on adult subjects, does not count against this view. To marshal the experience machine against this view, we must focus on subjects who are too young to have desires that are not merely experiential.

The relevant question, then, is whether Adam’s total welfare is higher during the neonatal period. I claim that it is. While Adam’s parents are playing with him, cuddling him, and taking care of him, Bill is being kept in total isolation and being fed a mere simulation of these events. No one loves Bill or cares about him, and he never interacts with another human being. Nor does he begin to develop his physical capacities or interact in any way with the real world. I claim that even if Bill’s total welfare during the neonatal period is nearly as high as Adam’s, it is at least a little bit lower:

Neonatal Period

Adam’s total well-being during the neonatal period is at least somewhat higher than Bill’s.

It is admittedly not as clear that Neonatal Period is true as it is that Adam is higher in total lifetime well-being. But this may be because the neonatal period is so short that the difference in welfare between Adam and Bill during this period is barely noticeable against the background of their total lifetime well-being. To get clearer on whether Neonatal Period is true, consider a tragic variant on the case, in which both babies die just before the end of the neonatal period. It is even clearer in the tragic variant that Adam’s total welfare during the neonatal period is greater than Bill’s: more welfare accrues to Adam during his tragically short life than does to Bill. Suppose you were entrusted with
the task of ensuring that a fetus who will soon be born and who will not survive beyond the neonatal period will have the best possible life. If you were ignoring any considerations except those relevant to its welfare, and you had to choose between giving it Adam’s short life or Bill’s, would you be totally indifferent between the two options? If not, then either you believe Neonatal Period or you believe that it might be true.

Neonatal Period is also supported by the observation that, while they are newborns, Bill is pitiful in a way that Adam is not. Intuitively, Bill’s condition warrants pity. But to pity someone is to feel bad for him on account of the fact that he is, in certain respects, badly off. If the pity that we feel toward Bill is warranted, then he really is in certain respects badly off. This suggests that Adam really is higher in welfare than Bill during the neonatal period.33

A Same World Desire Subjectivist might object: how could Adam be higher in welfare than Bill in virtue of a difference between their lives that neither of them cares about? But this would be question-begging: to ask this rhetorical question is just to ask how subjectivism could be false. Put aside your preexisting theoretical commitments and focus on the case at hand. Doesn’t it seem that Adam’s total welfare during the neonatal period is at least slightly higher than Bill’s?

If Neonatal Period is true, then Same World Desire Subjectivism is false. Recall that, unlike normal adults, newborns only have experiential desires. Thus, if neonatal Adam is higher in welfare than neonatal Bill, this cannot be because the profile of desires that they share is better satisfied in Adam’s case: their lives are experientially identical, so the desires that they have during this period are equally well satisfied. Nor could it be because the desires about the past that they will have later in life (e.g., a desire to have been raised by loving parents) are better satisfied during this period in Adam’s case. For in the tragic variant of the case, in which they both die when they are newborns, the neonatal difference in welfare persists. Since there are no future desires that could explain this difference in the tragic variant, it is surely not explained by future desires in the original case either. Thus, unlike the claim that Adam’s total lifetime well-being is higher than Bill’s, Neonatal Period cannot be explained by a difference in how well-satisfied Adam’s and Bill’s desires are. But

33 For a discussion of the relation between pity and well-being, see Campbell (2013), pp. 337, 351.
according to Same World Desire Subjectivism, only such a difference could account for a difference in their welfare. Thus, Neonatal Period rules out Same World Desire Subjectivism.

Notice that the assumption that newborns only have experiential desires is dispensable. Our inclination to accept Neonatal Period is not contingent on the supposition that, even when they are newborns, human beings have desires concerning the external world: we would feel no urge to revise our judgment that Adam is higher in welfare even if we obtained proof that newborns only have experiential desires (as I have assumed). This shows that my argument does not depend on the empirical claim that newborns only have experiential desires. It doesn’t matter whether newborns are actually like this. What matters is that Neonatal Period is true at possible worlds (actual or not) where Adam and Bill are as I described and where newborns only have experiential desires.

What if it were claimed that Same World Desire Subjectivism is merely intended to apply to normal adults? I would reply as I did earlier: the view cannot be false of newborns but true of normal adults unless there is some difference between them that could explain why this is so, and it is difficult to see what this might be. Indeed, it is even more difficult here than it was in the case of sophisticated Same World views, since both newborns and adults are capable of desire. Normal adults possess many capacities that newborns lack. But since none of these capacities is invoked by Same World Desire Subjectivism, it is not even prima facie plausible that this explains why the view might be true of normal adults even though it is false of newborns. Why would the view be true of one of these groups but not the other even though both groups possess the capacities that it considers necessary for well-being?

It might be suggested, for example, that Same World Desire Subjectivism becomes true of human beings when they become capable of having desires that are not merely experiential. Assuming, as I did, that newborns cannot have such desires, this would mean that Same World Desire Subjectivism becomes true of humans at some stage after the neonatal period. But the alleged fact that the view becomes true of us when we acquire this capacity cannot plausibly be explained. Same World Desire Subjectivism doesn’t say that the only desires whose objects are basically good are those that are not merely experiential. It doesn’t even accord greater weight to the objects of such desires than it does to the objects of experiential desires. Why, then, would the view nonetheless be true only of those subjects who can have desires that aren’t merely experiential? It seems that any other proposal on
which the view is true of normal adults in spite of being false of newborns would be committed to a similarly inexplicable divergence between neonatal and adult welfare. We should conclude that Same World Desire Subjectivism is false of newborns and normal adults alike.

Some readers might be inclined to resist my argument at an earlier stage by questioning Neonatal Period. Of course, theory choice is a holistic matter. If Neonatal Period is true, then Same World Desire Subjectivism is false. But Neonatal Period is not indubitable. Even if you have an intuition that it is true, your credence in it will depend on how impressed you are by the virtues of the view it rules out—and, more generally, on how willing you are to keep a simple theory in the running even when it has counterintuitive implications. Some philosophers greatly value simplicity and tend to mistrust intuitions about cases when they conflict with attractive, simple theories. Others put less weight on simplicity and more weight on fidelity to intuitions about cases. I side with the latter group, but it would be beyond the scope of this paper to show that they are correct.

5. Ideal World Subjectivism

I have argued against all three versions of Same World Subjectivism. But many theorists favor Ideal World Subjectivism, according to which something is basically good for you at \( W \) if and only if you have \( A \) toward it at \( I \), the closest world where you have full information (including acquaintance with all possible experiences) and full instrumental rationality. It might be thought that this view can evade my arguments. For presumably, if a newborn were fully informed, it would be sophisticated enough to value things and to have welfare beliefs, and it would have a profile of desires similar to that of a normal adult—one that includes desires that are not merely experiential. After all, a fully informed newborn would be cognitively and psychologically like an adult.

Many objections have been leveled against Ideal World Subjectivism. Some have argued that since having full information might cause you to have a bizarre or sparse profile of favorable attitudes,

\[ \text{References} \]

35 E.g., Nagel (1979, p. x), Temkin (2012, pp. 18-19), and Kamm (1992, pp. 1, 6-11).
36 I haven’t specified \textit{which} world \( I \) is closest to. The two most obvious possibilities are \( W \) and the actual world. Ideal World Subjectivists do not say much about this. I will assume that the relevant world is \( W \), but my arguments could straightforwardly be modified if it turned out to be the actual world.
this view leaves open many false hypotheses about your welfare: counting blades of grass might be the only thing toward which you have $A$ at $I$, but it surely isn’t the only thing that is basically good for you. 38 Others have claimed that this view accords too much influence to your actual personality: a rigid and overly cautious person might, when fully informed, retain these traits and therefore not want to become more carefree and spontaneous, but it doesn’t follow that it wouldn’t be good for her to change in these ways. 39 Moreover, different arguments have been given for the conclusion that it is impossible for anyone to be fully informed. 40 Since many philosophers remain sympathetic to Ideal World Subjectivism, I will assume for the sake of the argument that these objections can be answered. I will argue, however, that Ideal World Subjectivism should be rejected because it faces an additional problem when the subject is a newborn infant.

The problem is simple. Ideal World Subjectivism is a genuinely subjectivist view only if the following claim is true:

**Assumption**
Your personality at $I$ at least largely resembles your personality at $W$.

Indeed, Ideal World Subjectivism is a plausible view (subjectivist or not) only if that claim is true. But at least when the subject is a newborn, the Assumption is false. Thus, Ideal World Subjectivism is not a plausible view—or even a genuinely subjectivist view—about neonatal welfare.

Why is Ideal World Subjectivism a subjectivist view only if the Assumption is true? Recall that subjectivism says that something is basically good for you if and only if it *fits* or *suits* you. 41 This claim isn’t meant to apply only to the actual world: no subjectivist would say that if you had had a different upbringing, something could be basically good for you even though it didn’t suit you. Subjectivists claim that, at any possible world, what is basically good for you is what fits you *as you are at that world*. Now, this isn’t to say that $x$ is basically good for you at $W$ if and only if you have $A$ toward $x$ at $W$. That would be Same World Subjectivism, not subjectivism in general. But it is to say that, at any world $W$, what is basically good for you must ultimately (if somewhat circuitously) depend on what you are like at $W$, insofar as there must be a “match” or “fit” between your personality at $W$ and the

---

things that are basically good for you there. Even if what matters is not whether you have A toward x at W, but whether you have A toward x at some other world (e.g., I), that other world must be one where your personality is largely similar to the one you have at W. If you are a thrill-seeker, then the fact that you would like skydiving if you were a fully informed thrill-seeker might show that this activity suits you as you actually are. But if you are a risk-averse homebody, then the fact that you would like skydiving if you were a fully informed thrill-seeker in no way shows that this activity suits you as you actually are, and thus doesn’t show that it is basically good for you by subjectivist lights. Thus, if Ideal World Subjectivism is to be a genuinely subjectivist view, the Assumption must be true: I has to be a world at which the subject’s personality largely resembles his personality at W.\footnote{For a related discussion, see Loeb (1995), pp. 13-17.}

Why is Ideal World Subjectivism a \textit{plausible} view only if the Assumption is true? Because if the Assumption is false, the view implies that we cannot have as many justified beliefs about people’s welfare as we actually can. Given a sufficiently detailed description of someone’s personality and circumstances, we can arrive at many justified beliefs about her welfare. If we know that Jane is a thrill-seeker and an avid skydiver, for example, we are justified in believing that going skydiving would benefit her. According to Ideal World Subjectivism, this belief is justified only if we are justified in believing that Jane’s idealized self (Jane+) has a favorable attitude toward skydiving or things associated with it (e.g., thrilling experiences). If the Assumption is true, then we are justified in believing this, since we are justified in believing that Jane+ is, like Jane, a thrill-seeker. But if the Assumption is false, we cannot assume that Jane+ is also a thrill-seeker: for all we know, she is a risk-averse homebody who has no favorable attitudes toward any thrilling activities. So if Ideal World Subjectivism is true but the Assumption is false, fewer of our beliefs about people’s welfare are justified than actually are.\footnote{An Ideal World Subjectivist might propose some other function (besides resemblance) from a subject’s personality at W to his idealized self’s personality. But it is unclear what this function could plausibly be. For instance, it would clearly be implausible to suppose that one’s idealized personality is the opposite of one’s personality at W.}

The fact that Ideal World Subjectivism needs the Assumption has been recognized by defenders and critics of the view alike. For example, in response to the worry that we might have no idea what you desire at I, Railton says that your personality at I is the result of changing your personality at W in the “entirely familiar” ways in which we know people to change in response to new information or experiences. Indeed, he adds, we hold the subject’s personality “as nearly constant as possible when
asking what someone like him would come to desire” if fully informed and rational. And Rosati, a critic of Ideal World Subjectivism, attributes the Assumption to it: “while this ideal standpoint [viz., I] alters a person’s epistemic condition, it does not by design alter her personal characteristics…. On the contrary, the proposed idealization holds a person’s personality constant, permitting only those changes that would result from fully informing her and improving her reasoning.44

I am assuming for the sake of the argument that the Assumption is true when the welfare subject is an adult—or, at least, that it can be made true via suitable refinements to how the ideal world, I, is specified. But the Assumption cannot be true when the subject is a newborn infant. If a newborn were fully informed and rational, he would have a personality at least as complex and well-developed as that of a typical adult: for if he knew everything and had firsthand acquaintance with all possible experiences, he would be a seasoned veteran in the art of living, not a neophyte. The personality of a subject who is psychologically like a superhuman adult cannot largely resemble that of an infant who has just emerged from the womb.

This means that, as a view about the welfare of newborns, Ideal World Subjectivism is not genuinely subjectivist: it does not honor the thought that what is basically good for someone at W is what suits her as she is at W’. It also means that, as a view about newborns, this view is implausible: for it implies that we cannot have as many justified beliefs about a newborn’s welfare as we actually can. Since these problems do not depend on any particular way of specifying attitude A, they affect all versions of Ideal World Subjectivism. The first problem gives subjectivists sufficient reason to conclude that no version of this view is true of newborns. The second one gives everyone some reason to conclude this.46

It might be objected that the Assumption needn’t be accepted by Ideal World Subjectivists like Railton, who claim that the relevant attitudes are the ones you have at I toward your wanting,

44 Railton (2003b), pp. 57, 60.
46 Rosati (2009, pp. 208-09) briefly argues that full information theories such as Railton’s cannot be true of infants. In her view, the problem is that what an infant would desire if fully informed depends on what motivational system she would have if fully informed. “Baby might well develop in any number of ways, resulting in different Baby pluses with possibly differing desires for actual Baby. There seems no single, relevant counterfactual, and so no determinate question as to what Baby+ would want to be done or pursued for Baby.” But this doesn’t strike me as decisive. After all, even when the subject is an adult, there are many fully informed Adults with different personalities. This is no problem: we just look at the fully informed Adult who is at the closest world to W’ out of all of the worlds where he is fully informed.
pursuing, or having $x$ at $W$. In their view, what matters isn’t whether Jane+ wants to engage in thrill-seeking activities, but whether she wants Jane to do so. And presumably, Jane+ wants Jane to engage in thrill-seeking activities whether or not Jane+ wants to engage in such activities herself: for Jane+ knows what Jane wants, and presumably, she wants Jane to have exactly what Jane wants. Thus, on such views, it doesn’t matter how dissimilar a subject’s personality is from that of her idealized self: the idealized self’s attitudes regarding what the subject wants, pursues, or gets at $W$ will simply defer to the subject’s own attitudes at $W$. So even if the Assumption is false, such views are subjectivist: they guarantee that what’s basically good for Jane at $W$ is what suits her at $W$, since they say that Jane+’s favorable attitudes regarding what Jane gets at $W$ perfectly mirror Jane’s own favorable attitudes at $W$. And even if the Assumption is false, such views don’t imply that we cannot have as many justified beliefs about people’s welfare as we actually can: since Jane+’s attitudes concerning Jane mirror Jane’s own attitudes, a sufficiently detailed description of Jane’s attitudes and circumstances will enable us to work out how well off she is.

This objection relies on a dubious reading of views like Railton’s, however. If such views maintain that a subject’s idealized self always defers to the subject’s own attitudes, then they are really just disguised versions of Same World Subjectivism, and they are vulnerable to my earlier arguments. (If $x$ is basically good for Jane at $W$ if and only if Jane+ wants Jane to have it at $W$, and Jane+ wants Jane to have $x$ at $W$ if and only if Jane wants it at $W$, then $x$ is basically good for Jane at $W$ if and only if Jane wants it at $W$.) Thus, in the present context, we should interpret such views as holding that Jane+’s attitudes concerning Jane might not mirror Jane’s own attitudes. But this brings us back to the importance of the Assumption. If Jane+’s personality doesn’t largely resemble Jane’s, and Jane+ doesn’t want for Jane whatever Jane wants for herself, then there is no guarantee that what Jane+ wants for Jane suits Jane. Indeed, it is hard to know what Jane+ wants for Jane, and thus hard to know what is good for Jane on these views. Thus, if views like Railton’s are not merely disguised versions of Same World Subjectivism, they require the truth of the Assumption. But as I argued above, the Assumption is false when the subject is a newborn. Thus, my point stands: Ideal World Subjectivism is implausible, and not genuinely subjectivist, when applied to newborns.

If, as any subjectivist must acknowledge, Ideal World Subjectivism is false of newborns, we should presume that it is also false of normal adults. For if it were true of normal adults even though it is false of newborns, there would have to be some difference between them that explains this fact. I
see no such difference. If nothing about the favorable attitudes of your fully informed self explains why anything is basically good for you when you are a newborn, why would reaching a certain stage of development suddenly make it true that everything that is basically good for you has that status because you would have the right favorable attitude toward it if you were fully informed? Absent an answer to this question, we should conclude that Ideal World Subjectivism is false of adults.

It is unlikely that this question can credibly be answered. To see why, imagine a human being who is high in welfare and who is approaching the time at which Ideal World Subjectivism allegedly becomes true of her. It could be that some of the things that are basically good for her are such that her idealized self doesn’t have the right kind of favorable attitude toward them. (Even the Ideal World Subjectivist would agree, since he concedes that his view is not yet true of her.) Moreover, it could surely remain true, even after she passes the point at which Ideal World Subjectivism allegedly becomes true of her, that her idealized self lacks the right kind of favorable attitude toward those things. But if Ideal World Subjectivism becomes true of her at that point, then from then onward, those things fail to be basically good for her—even though the facts about her idealized self’s attitudes towards them have remained completely unchanged. This is implausible. Thus, since we should reject Ideal World Subjectivism for newborns, we should likewise reject it for normal adults.47

6. Objections and Replies

6.1 Different Theories Might Be True of Newborns and Normal Adults

It might be objected that, since newborns are arguably as different from normal adults as dogs and other animals are, we shouldn’t assume that the correct theory of their welfare is the same as the correct theory of the welfare of adults. Indeed, it might even be claimed that we should assume that different theories are true of those two groups. But if this is so, then it seems that, contrary to what I have argued, we shouldn’t conclude that the subjectivist views I have discussed are false of normal adults on the grounds they are false of newborns.48

47 My argument in this section also counts against a weaker version of Ideal World Subjectivism—one that merely says that x is basically good for you only if (and at least partly because) you have A toward it at I. It is not plausible (or truly subjectivist) to say that x is basically good for you only if it is favored by a version of you who might be nothing like you.

48 I thank an anonymous referee for raising this objection.
This objection assumes that my arguments have relied on the principle that the true theory of the welfare of normal adults, \textit{whatever} it turns out to be, is also the true theory of neonatal welfare. Since I have argued that each of the subjectivist views I have discussed is false of normal adults because it is false of newborns, it would be natural enough to suppose that I rely on this principle. I am, in fact, inclined to accept this principle; and elsewhere, I have tried to explain why, contrary to appearances, the many differences between normal adults and newborns do not undermine it.\footnote{Lin (unpublished b). Indeed, in that paper, I suggest that the same theory of welfare is true of all welfare subjects, human or not. But nowhere in the present paper do I assume this view.} However, I have \textit{not} relied on this principle in this paper. I did not argue that none of the aforementioned subjectivist theories can be true of normal adults on the general grounds that \textit{no} theory of welfare can be true of normal adults but false of newborns. Rather, for each of those subjectivist theories, I argued that it is implausible that \textit{that very theory} could be true of normal adults even though it is false of newborns. For example, I argued that Same World Judgment Subjectivism is false of normal adults for the following reason: if it were true of normal adults even though it is false of newborns, then someone could (implausibly) go from being high in welfare to being neutral or negative in welfare simply in virtue of acquiring the capacity to have welfare beliefs. This does not assume that \textit{no} theory could plausibly be true of normal adults even though it is false of newborns. For all that I have said, there are such theories: I merely claim that the aforementioned subjectivist theories are not among them. Thus, my arguments do not depend on the view that whatever theory is true of normal adults must also be true of newborns. For this reason, they evade the present objection.

Admittedly, I endorsed Difference, which implies that a theory of welfare cannot be true of one kind of subject but false of another unless there is some difference between the two kinds that explains why this is so. But that claim is relatively uncontroversial: those who think that different theories are true of different kinds of subject would surely agree that certain differences between those kinds explain this. I have merely argued that no difference between normal adults and newborns can explain why any of \textit{the particular views that I have discussed} might be true of the former but false of the latter. Thus, you can accept my arguments even if you believe that different theories are true of the two kinds of subject.

There may be a different objection in the vicinity of this one, however. It might be objected that I haven’t \textit{established} that no plausible explanation could be given for why any of the subjectivist views I
discussed might be true of normal adults even though it is false of newborns. For example, I haven’t decisively ruled out the possibility that someone might produce a plausible explanation of why someone could go from being very high in welfare to being neutral (or negative) in welfare simply in virtue of acquiring the capacity to have welfare beliefs—an explanation in light of which it might no longer seem so implausible that this could happen. I concede that this is true, but I don’t think it significantly undermines the importance of my arguments. I have given new reasons to reject several subjectivist views. I believe that they are strong reasons. But they are, in principle, defeasible. In this respect, my arguments are no different from most other arguments in philosophy.

6.2 The Threat of Overgeneralization

It might be wondered whether the problems that I have raised would also apply to other theories of welfare, and in particular, to non-subjectivist theories such as objective list theories. If so, then since the subjectivist views I have targeted may be no worse than some of their main competitors, I haven’t provided very strong reasons to reject those views in favor of those competitors. Indeed, perhaps the problems I have raised generalize to too many theories: if my arguments leave us with no plausible candidates for the correct theory of our welfare, then there is likely something wrong with them.\textsuperscript{50}

This is an important objection. But fortunately, the problems do not overgeneralize. Consider an objective list theory whose list of basic goods includes pleasure, knowledge, moral virtue, and being loved.\textsuperscript{51} Since newborns are capable of pleasure and of being loved, this theory doesn’t imply that nothing is basically good for them, so it avoids the problem faced by sophisticated Same World subjectivist views. Since this theory doesn’t idealize, it evades the problem faced by Ideal World subjectivist views. Finally, this theory avoids the problem faced by Same World Desire Subjectivism: it says that since Adam is loved but Bill is not during the neonatal period, Adam is at least slightly higher in welfare during that period. This theory is not subjectivist, since it denies that the basic prudential value of a particular thing is wholly fixed by the extent to which you have a favorable

\textsuperscript{50} I thank an anonymous referee for raising this worry.

\textsuperscript{51} Recall (from Section 1) that basic goods are basically good \textit{kinds}: they are to be distinguished from their members, which are basically good \textit{particulars}. 
attitude toward it under the right conditions. It is plausible. So it is a plausible, non-subjectivist view to which my problems do not apply.

Someone might object that this theory faces a similar problem to the one faced by sophisticated Same World subjectivist views. Knowledge and moral virtue are inaccessible to newborns, in the sense that newborns cannot possess, or otherwise be appropriately related to, tokens of these kinds. Doesn’t it follow that they are not basic goods for newborns, and thus that this theory is false of newborns?

This line of reasoning is extremely natural, but I believe it relies on a mistaken view about what it takes for a certain kind (e.g., moral virtue) to be a basic good for a given kind of welfare subject. Consider a cognitively normal adult who never instantiates any moral virtue during his lifetime. Indeed, imagine that his failure to instantiate virtue is fairly modally robust: for a great many of the ways that things could go, he would still never instantiate virtue if things were to go that way. Does the possibility of such a person show that our objective list theory is not true of all normal adults? No. For the claim that virtue is a basic good for this subject doesn’t imply that he actually possesses any tokens of virtue. Nor does it even imply that he possesses some tokens of virtue at any nearby possible worlds. It merely implies this: for any token of virtue, if he were to possess or instantiate that token, it would be basically good for him. In spite of the possibility of a person so thoroughly lacking in virtue, our objective list theory could be true of all normal adults. For the following counterfactual could be true of every normal adult: for any token of virtue, if he were to possess that token, it would be basically good for him. This discussion shows that the claim that $K$ is a basic good for subjects of a certain kind merely implies that, for any token of $K$ and any subject $S$ of that kind, if $S$ were to possess that token, it would be basically good for $S$.\footnote{I defend this view at greater length in Lin (unpublished b).}

With this view in hand, we can see why the fact that knowledge and virtue are inaccessible to newborns does not imply that they are not basic goods for newborns. Newborns possess no tokens of knowledge or virtue hereabouts in modal space. But it could nonetheless be true that if any newborn were to possess any token of either of those kinds, that token would be basically good for it.
Thus, the inaccessibility to newborns of knowledge and virtue doesn’t show that our objective list theory is false of newborns.

I should emphasize that these remarks do not undermine my earlier arguments that sophisticated Same World subjectivist views are false of newborns. For all that I have said, if any newborn were to believe anything to be basically good for him, it would be basically good for him. Thus, for all I have said, the things one believes to be basically good for one are a basic good for newborns. But this doesn’t show that Same World Judgment Subjectivism could be true of newborns. For this view says not merely that the things you believe to be basically good for you are a basic good, but that they are the only one. Thus, it implies that since newborns actually fail to believe anything to be good for them, no particulars are actually basically good for them. It is on the basis of this false implication—not the mere fact that newborns actually lack welfare beliefs—that I concluded that this view is false of newborns. By contrast, our objective list theory does not imply that no particulars are actually basically good for newborns. After all, it says that pleasure and being loved are both basic goods, and newborns can actually possess tokens of both these kinds.

If, as I have argued, the aforementioned objective list theory can be true of newborns, then we needn’t conclude that it is false of normal adults because it is false of newborns. Thus, my arguments leave it open that this theory (and others relevantly like it) might be the correct theory of our welfare. Indeed, even if we did have to think that this theory is false of newborns, none of my arguments implies that we would therefore have to think that it is false of us. As I said earlier, I do not claim in this paper that the same theory of welfare must be true of newborns and normal adults. For all that I have said, some difference between the two groups could explain why our objective list theory is true of normal adults even if it is false of newborns. If, contrary to what I have just argued, the fact that knowledge and virtue are inaccessible to newborns does entail that this theory is false of them, then this very fact could explain why the theory might nonetheless be true of us.

It bears emphasizing that, since our objective list theory can be true of newborns and normal adults alike, it has nothing like the implausible implications that sophisticated Same World subjectivist views have about cases in which someone who is already high in welfare acquires certain welfare-relevant capacities but fails to exercise them. Since the latter views are false of subjects who lack the capacities they invoke, they imply that in such cases, the subject’s welfare goes from high to zero (or
negative) simply because he acquires the relevant capacity. But the objective list theory doesn’t imply this. Imagine a newborn who is high in welfare because he is being loved to a certain extent and has a certain favorable balance of pleasure over pain. Suppose that he acquires the capacities to have knowledge and moral virtue, but that he does not immediately exercise them. The theory does not imply that his welfare drops precipitously at this time: if nothing else has changed, then he retains the same level of welfare that he previously had (though he would be even better off if he were to instantiate some knowledge or virtue).

My arguments do apply to some objective list theories, such as the view that the only basic goods are virtue, knowledge, and pleasure taken in intellectual contemplation. Since newborns cannot possess tokens of these kinds, this view falsely implies that nothing is basically good for them. Thus, at least one other kind must be a basic good for newborns. But if this view is true of normal adults, this other kind inexplicably loses its status as a basic good when a human reaches the stage of development at which the view becomes true of him. We should therefore reject this view even as a theory of the welfare of normal adults. The fact that my arguments count against some objective list theories is no problem, however: as I have argued, other such theories can withstand those arguments. It is an advantage of my arguments that they count against objective list theories that require too much cognitive sophistication on the part of the subject.

I should also mention that my arguments leave open certain non-subjectivist views that are not objective list theories. For instance, consider a “hybrid” view on which something is basically good for you if and only if you desire it, but how basically good for you it is depends both on how strongly you desire it and the extent to which it is worthy of desire. Since newborns are capable of desire, this view avoids the problem that I raised for sophisticated Same World views. Since this view doesn’t idealize, it is immune to the problem that I raised for Ideal World Subjectivism. Finally, assuming that experiences that are caused by interactions with the real world (or other people) are more worthy of desire than ones caused by an experience machine, this view also says the intuitively right thing about Adam and Bill. Thus, it avoids the problem faced by Same World Desire Subjectivism.

As I hope the foregoing remarks have shown, my arguments do not generalize to too many theories of welfare, or to too many non-subjectivist ones: they leave us with a number of plausible theories of the welfare of normal human adults, including some objective list theories and hybrid theories.
7. Conclusion

The six views that I have discussed do not exhaust all possible subjectivist views. However, practically every extant subjectivist view is one of these views or a hybrid thereof. Even views that don’t fit neatly into my taxonomy are vulnerable to my arguments.

Sumner, for instance, claims that you are well off to the extent that you are happy, provided that your happiness is authentic. Happiness “has both an affective component (experiencing the conditions of your life as fulfilling or rewarding) and a cognitive component (judging that your life is going well for you, by your standards for it).” Happiness is authentic just if it is informed and autonomous. Thus, Sumnerian Subjectivism says that \( x \) is basically good for you at \( W \) if and only if at \( W \), you authentically (i) judge that \( x \) is basically good for you and (ii) experience \( x \) as rewarding or fulfilling.\(^53\) As we’ve already seen, newborns do not have welfare beliefs, so they cannot satisfy (i). Nor can they satisfy (ii). For as Sumner emphasizes, to experience something as rewarding or fulfilling—or, as he puts it elsewhere, “enriching,” “satisfying,” or “worthwhile”—is not merely to take pleasure in it or to enjoy it.\(^54\) Newborns can enjoy things and perhaps even feel joy, but they are not sophisticated enough to experience anything as enriching, fulfilling, or worthwhile. Thus, Sumner’s view is too sophisticated twice over.\(^55, 56\)

The same is true of the Values-Based Life Satisfaction view, defended by Valerie Tiberius and Alexandra Plakias.\(^57\) On this view, something is basically good for you only if you judge that it meets certain standards set by your values—where valuing something involves being stably disposed to have a positive emotional response to it and to take it as providing you with reasons for doing certain things. This view can’t be true of newborns, since they can’t have values in this sense, and they can’t judge that anything meets certain standards.

\(^{53}\) Sumner (1996), p. 172. At one point (pp. 172-73), Sumner writes as though conditions (i) and (ii) should be disjoined rather than conjoined. But he has confirmed via personal communication that this passage is misleading.

\(^{54}\) Sumner (1996), pp. 146-49. For a similar point about fulfillment, see Wolf (1997), pp. 216-18.

\(^{55}\) At one point, Sumner claims that according to his theory, all sentient creatures are capable of welfare (1996, p. 178). I don’t see how this could be true. A creature could be capable of feeling pleasure or pain without being capable of happiness, as he describes it. Sumner is sensitive to the fact that a theory of welfare shouldn’t imply that the range of welfare subjects is narrower than it is (1996, pp. 14, 178). But it seems to me that this is exactly what his theory does.

\(^{56}\) Rosati (2009, pp. 207-08) claims that Sumner’s view applies “arguably only to normal and nearly adult humans.”

\(^{57}\) Tiberius and Plakias (2010).
Of course, it might be claimed that these views are true of subjects who possess the relevant capacities (e.g., normal adults) even though they are false of newborns. But I would reply as I did earlier: if this were true, then someone could go from being high in welfare to being neutral or negative in welfare simply in virtue of acquiring the relevant capacities. This is implausible, so we should think that these views are false even of adults.

I have argued that nearly all extant subjectivist views face a dilemma: either they have implausible implications about the welfare of newborns, or they exclude newborns from their scope but cannot plausibly explain this exclusion. My discussion also yields a general lesson: when theorizing about well-being, it is a mistake to focus only on a small subset of the entities that are capable of well-being. If you do this, you run the risk of endorsing claims that are implausible because they imply that the facts about the welfare of those entities diverge from the facts about the welfare of other entities in ways that cannot plausibly be explained.

Skelton (2015) interprets Sumner as accepting something like hedonism for young children. In this paper, I have focused on monistic subjectivist views—subjectivist views on which there is only one basic good. However, as I explain in Lin (forthcoming b), pluralistic subjectivist views—or as I call them, subjective list theories—are also possible. My arguments, or variants thereof, would also count against many such theories. For example, my argument against Same World Desire Subjectivism would also apply to Disjunctive Desire Satisfactionism, the view that there are two basic goods: desire satisfaction and subjective desire satisfaction.

I thank Katalin Balog, Nick Beckstead, Stephen Campbell, Richard Yetter Chappell, Dale Dorsey, David Frank, Josh Gillon, Chris Heathwood, Sari Kisilevsky, Sarah-Jane Leslie, Barry Maguire, Sarah McGrath, Jennifer Morton, Gideon Rosen, Michael Smith, Gerard Vong, and several anonymous referees.
Works Cited


_____ (unpublished), “Which Desires Are Relevant to Well-Being?”


