

# Subjective Theories of Ill-Being<sup>1</sup>

Anthony Kelley  
Louisiana State University  
akelley@lsu.edu

~12.5K words, including footnotes and bibliography. Forthcoming in *Midwest Studies in Philosophy*.  
Cite published version.

Philosophers working on welfare have historically focused on *well-being* and what is *good* for you. Recently, at the invitation of Shelly Kagan, we have turned our attention to the negative side of welfare concerning *ill-being* and what is *bad* for you. This paper contributes to the growing literature on this topic. My interest here is in extending the popular, yet controversial subjectivist tradition concerning well-being to cover the new and burgeoning field of ill-being. According to subjective theories of ill-being, the token states of affairs that are basically bad for you must be suitably connected, under the proper conditions, to your negative attitudes. This paper explores the prospects for this family of theories and addresses some of its challenges. My main aims are (i) to show that subjectivism about ill-being can be derived from a more general doctrine that requires a negatively valenced relationship between any welfare subject and the token states that are of basic harm to that subject and (ii) to respond to some objections, including the objection that subjectivists about ill-being cannot plausibly explain the badness of pain.

## I. Introduction

According to subjectivism about well-being, the token states of affairs that are basically good for you must be suitably connected, under the proper conditions, to your positive attitudes. Perhaps the most prominent subjective theory is desire satisfactionism, or the desire-fulfillment theory, according to which you are better off to the extent that you get what you want.<sup>2</sup> The theory is subjective because desire is a positive attitude, and the theory guarantees that each thing that is basically good for you is suitably connected to a desire of yours. Desire theorists, like proponents of other subjective theories, disagree over whether the proper conditions should be specified as the conditions of the world in which the subject exists or the conditions of some other possible world, such as one in which the subject is idealized in one or more respects. Subjectivists also disagree

---

<sup>1</sup> For helpful discussion, I thank Teresa Bruno-Niño, Dale Dorsey, Chris Heathwood, Eden Lin, Joseph Van Weelden, and the students in my fall 2022 well-being seminar. For written feedback, I thank Gwen Bradford.

<sup>2</sup> The desire theory is widely discussed in the literature. For a representative sample, see Sidgwick 1907 (I.IX.3), Brandt 1979, Rawls 1971, Singer 1979, Parfit 1984 (Appendix I), and Heathwood 2016.

about which connections count as suitable, but a standard way of proceeding is to require that the basic good state is the object of the relevant positive attitude(s).<sup>3</sup> Subjectivists typically pick a single positive attitude whose satisfaction or fulfillment is necessary for benefit, though it is possible to construct a subjective theory that countenances a plurality of positive attitudes as welfare-relevant. We can get different kinds of subjective theories by altering the number or kind of positive attitudes to which the theory makes reference.

As a class, subjective theories derive much of their plausibility from a certain doctrine concerning the nature of prudential value. The doctrine in question enjoys no uniform name or formulation. According to what I will refer to as the *resonance constraint*, there must be a positive connection, or a resonance, between a welfare subject and each token state that is basically good for that subject. This positive connection need not be univocal across types of subjects; how it is realized for a particular subject may depend on that subject's nature. For beings like us with the capacities that we have, the positive connection is realized through our evaluative attitudes, such as our desires and our evaluative beliefs. There is widespread agreement that subjective theories respect the resonance constraint since they guarantee that the token states of basic benefit to you are connected to your positive attitudes, and this connection is taken to be sufficient for the kind of positive connection referenced in the resonance constraint.<sup>4</sup> Subjectivism about well-being rules out a wide-range of competing theories, including perfectionism and objective-list theories that posit

---

<sup>3</sup> Another option would be to regard the suitable connection more broadly. On a typical combo subjectivist view, the basically good state is the combination state of affairs consisting in the relevant positive attitude and its object. According to such views, the basically good state is thought to be suitably connected to a positive attitude of yours, not by being the object of that attitude, but instead by constitutively involving it. For a discussion (and rejection) of the combo version of the desire theory, see Van Weelden 2019.

<sup>4</sup> I have argued elsewhere, however, that this connection is *not* sufficient for the kind of positive connection referenced in the resonance constraint. See Kelley 2020 (chapter one).

goods such as knowledge, achievement, or loving relationships without requiring that those goods be connected, in the intended sense, to the subject's positive attitudes.<sup>5</sup>

Everything that has been said so far has to do with the positive side of welfare concerning *well-being* and welfare *goods*. This paper, however, has to do with the negative side of welfare concerning *ill-being* and welfare *bads*. I began by describing subjectivism about well-being because my aim in this paper is to explore the prospects of an analogous family of theories about ill-being:

**Subjectivism About Ill-Being:** A token state of affairs,  $x$ , is basically bad for a welfare subject who is capable of having negative attitudes only if and (at least in part) because  $x$  is suitably connected, under the proper conditions, to that subject's negative attitudes.<sup>6</sup>

Much of what has just been said about subjectivism about well-being can be said, *mutatis mutandis*, about subjectivism about ill-being. Subjectivists about ill-being will disagree about how to specify the proper-conditions clause, which negative attitudes are relevant, and which connections count as suitable. I assume for the purposes of this paper an ecumenical stance with respect to which negative attitudes are relevant since my purpose here is to defend subjective theories of ill-being as a class, rather than any particular subjective theory. For ease of reference, I also assume that a state of affairs is suitably connected to the relevant negative attitude if and only if it is the object of that attitude. Lastly, I will sometimes drop the proper-conditions clause when stating either subjectivism about

---

<sup>5</sup> Some objective-list theorists argue that their view can accommodate the resonance constraint, provided that the list of welfare goods only includes items that constitutively involve positive attitudes, such as friendship and certain other interpersonal relationships. See Fletcher 2013.

<sup>6</sup> Others may wish to characterize subjectivism about ill-being more narrowly as the doctrine that  $x$ 's suitable connection, under the proper conditions, to  $S$ 's negative attitudes is both necessary *and* sufficient for  $x$  to be basically bad for  $S$ . My interest here is in defending the broader class of subjective theories as defined in the main text. A hybrid view according to which  $x$  is basically bad for you only if you have a negative attitude towards  $x$  and  $x$  is impersonally objectively bad would count here as a subjective theory. See Kagan 2009. Note that such a theory would count as subjective in my sense even if the degree to which  $x$  is bad for you supervenes on  $x$ 's degree of objective badness and not (just) the strength of the negative attitude. For discussion of such object sensitivity in hybrid theories of ill-being, see Kagan 2021 (pp. 128-129).

well-being or subjectivism about ill-being, but readers should understand it to be implicit in the discussion that follows.

Just as subjectivism about well-being derives much of its plausibility from the resonance constraint, subjectivism about ill-being can be derived from what I call the *dissonance constraint*, the doctrine that there must be a negative connection, or a dissonance, between a welfare subject and each token state that is basically bad for that subject.<sup>7</sup> Subjectivism about ill-being would rule out objective theories of ill-being that claim, for example, that states of ignorance are basically bad for you, irrespective of how they are connected to your negative attitudes.

Philosophers have only recently turned their attention to developing and evaluating theories of ill-being, and this paper contributes to this growing literature.<sup>8</sup> Theories of ill-being are crucial for constructing a comprehensive theory of welfare that can tell us, together with the relevant descriptive facts at the relevant possible worlds, how a subject is faring at a time and during an interval of time. Any plausible moral theory will say that an act's moral status is at least partly determined by how it affects the welfare of individuals. Additionally, policy makers must take into account a policy proposal's potential impact on individuals' welfare when deciding whether to implement it. The present project is thus of interest to philosophers working in value theory and the normative ethics of behavior, as well as policy makers interested in theoretically informed decision making.

---

<sup>7</sup> Subjectivism about ill-being should be understood as applying only to those welfare subjects who are capable of negative attitudes. The dissonance constraint is compatible with the existence of welfare subjects incapable of such attitudes. For those subjects, there would be some other kind of negative connection between them and the states that are basically bad for them.

<sup>8</sup> In addition to the papers included in this issue of this journal, some recent philosophical work on ill-being includes Kagan 2014, Tully 2017, Mathison 2018, Rice 2019, Sumner 2020, Bradford 2020, Kelley 2020, Bradford 2021, and Kagan 2021.

## II. Preliminaries

Theories of welfare tell us how a person is faring at a time and during an interval of time, including the interval of time that encompasses the person's life as a whole. A comprehensive theory of welfare is constituted by both a theory of well-being and a theory of ill-being. Whereas a theory of well-being specifies the basic goods, a theory of ill-being specifies the basic bads. I will classify as subjective any theory of ill-being that implies subjectivism about ill-being, and I will regard a subjective theory of welfare as one that includes both a subjective theory of well-being and a subjective theory of ill-being.

Our concern is with states of affairs that are of *basic* value or disvalue.<sup>9</sup> These are the states that have value or disvalue for a subject non-derivatively and not just in virtue of how they are related to something else of value, such as being its cause or being composed of it. Theories of welfare specify the basic good and bad *kinds* whose instances are good or bad for you, respectively. For example, hedonism says the basic good kind is *your experiencing pleasure* and that each instance of this kind—that is, each pleasurable experience of yours—is basically good for you precisely because it is an instance of the basic good kind specified by the theory.<sup>10</sup> A theory of welfare will also specify the degree of goodness and the degree of badness of each state, such that, to the extent possible, for every possibly obtaining state of affairs, the theory would yield a rank-ordered list of states that would benefit and states that would harm the subject, ordered according to the degree to which they would be beneficial and the degree to which they would be harmful, respectively.

It is natural to expect a kind of symmetry in a theory of welfare between the theory of well-being and the theory of ill-being that compose it. If a theory says that  $x$  is a basic good for you,

---

<sup>9</sup> I will at times drop the 'basic' qualifier for stylistic reasons in the discussion that follows, but all references to value should be understood as references to basic value unless otherwise specified.

<sup>10</sup> For more on this way of understanding theories of welfare, see Lin 2017a.

then it is plausible that its opposite is basically bad for you.<sup>11</sup> Hedonism is a theory that exemplifies this kind of symmetry since the theory says that pleasure is good for you and its opposite, pain, is bad for you.<sup>12</sup> Subjectivists about welfare seem to have an advantage with respect to honoring the symmetry because they appeal to the attitudes of the subject, and attitudes typically have relatively easily identifiable pairs of polar opposites (e.g., desire and aversion, valuing and devaluing). Objectivists, on the other hand, make no such appeal to the attitudes of the subject, so they cannot avail themselves of these resources to explain the alleged symmetry. Indeed, objectivists have difficulty specifying what can even be plausibly thought to be the opposite of complex states such as knowledge, which many objectivists claim is a basic welfare good.<sup>13</sup>

Our topic is ill-being. Shelly Kagan argues that a theory of ill-being must postulate what he calls *robust bads*, which are states of affairs that “directly constitute a life’s going badly.”<sup>14</sup> I take it that robust bads are the things that are basically bad for you, so they are the states of affairs specified by a theory of ill-being. To illustrate the concept of a robust bad, Kagan points out that the hedonist

---

<sup>11</sup> It is possible that some goods have more than one opposite. If so, then we can distinguish between two importantly different claims of symmetry. First, there is the claim that there is a basic bad kind that is the opposite of each basic good kind. Second, there is the claim that there is a *unique* basic bad kind that is the opposite of each basic good kind. To see the difference, suppose that *G* is a basic good kind according to some theory of welfare and that *X*, *Y*, and *Z* are the opposites of *G*. If the theory says that *X*, *Y*, and *Z* are all basic bad kinds, then the theory respects the first symmetry claim but not the second because though the basic bad kinds it identifies are each opposites of the basic good kind identified by the theory, the theory does not say that there is a single, unique basic bad kind that is the opposite of *G*. Kagan (2014, p. 277) briefly discusses this distinction.

<sup>12</sup> It is not uncontroversial that pain is the opposite of pleasure. Some argue that the opposite of pleasure is the broader category of the unpleasant, such as feelings of dizziness or nausea that, allegedly, fall short of being painful. For discussion, see Heathwood 2007 (section 6.5).

<sup>13</sup> There is another kind of symmetry that we might expect between a theory of well-being and a theory of ill-being. For example, in addition to each basic good kind having an opposite basic bad kind, we might also think that the degrees of goodness and the degrees of badness of comparable goods and bads must be symmetrical. For example, a version of hedonism according to which a quantity of pleasure counts just as much for well-being as the same quantity of pain counts for ill-being is symmetrical in this sense. A different kind of view would say that a quantity of pain makes you worse off to a greater degree than the degree to which a comparable quantity of pleasure makes you better off. This latter view would respect the kind of symmetry I mention in the main text since it posits an opposite basic bad kind for each basic good kind, but it violates this second kind of symmetry since the basic bad particulars are more harmful than comparable basic good particulars are beneficial. For discussion, see Mathison 2018 (chapters one and two).

<sup>14</sup> Kagan 2014 (p. 262).

says that pain—a state that is the presence of a bad state rather than the mere absence of a good one—is basically bad for you. Pain satisfies Kagan as a plausible candidate for a robust bad. It is not clear, however, how to best characterize the concept of a *robust bad*. On some occasions, Kagan describes robust bads as states that “directly contribute to a life going badly,” and on others he describes them as states that “directly constitute a life’s going badly.”<sup>15</sup> It is not obvious that these are just different ways of describing the same concept. If we understand ‘contributing’ as *making more likely to obtain*, states that *contribute* to a life going badly cannot themselves *constitute* a life going badly, on the plausible assumption that no state of affairs contributes to its own obtaining.

Furthermore, we may wish to avoid a conceptual link between robust bads and lives going badly. Even if it were not possible to live a bad life, it would seem that something could still be robustly bad for you. Whether a life is overall bad for the person who lives it is simply a matter of how we weigh the goods and the bads present in that life. Even if we had an admittedly peculiar way of calculating the overall goodness of a life according to which no life is bad for the person who lives it, it seems that we would not have thereby ruled out the possibility of robust bads. Moreover, the badness of a life consists in not just the presence of bad things in that life. It is possible, for example, for two lives to contain the same number and degree of basic bads and for only one of them to count as a bad life for the person who lived it. This could happen when the basic bads in one of the lives are outweighed by the basic goods in that life. Since facts about the quality of a subject’s life consist in facts about the basic bads *and* the basic goods that accrue to the subject, it is too broad of a characterization to say that robust bads are those states that constitute a life’s going badly.

To avoid this conceptual link between robust bads and lives going badly we might want to say that a robust bad is a state that “directly constitutes a reduction in well-being,” which is yet

---

<sup>15</sup> Ibid. (pp. 261-262).

another way that Kagan characterizes the target concept.<sup>16</sup> This characterization is also faulty. Suppose, for example, that you are no longer enjoying a pleasure that you were enjoying just a moment ago. Suppose further that all else is equal in terms of the goods and bads in your life such that you are now faring on the whole just a bit less well now that you are no longer experiencing that pleasure. The mere absence of pleasure in this case constitutes a reduction in well-being and would count as a robust bad according to the proposal currently under consideration. But Kagan regards the mere absence of pleasure as the paradigmatic case of a thing that does not count as a robust bad, so directly constituting a reduction in well-being is apparently insufficient to count as a robust bad.<sup>17</sup>

Even if we were able to successfully analyze the concept of *robust bads*, one might wonder whether we need a theory of ill-being at all. Why should we think that there are any robust bads? Perhaps there are only goods to be attained but no robust bads, *per se*. Kagan says that views which deny the existence of robust bads are “incredible, implausible” and seemingly “clearly false,” but he offers no argument in support of these claims.<sup>18</sup> There is, however, a compelling line of reasoning at the ready. Begin with the thought that it seems possible to have a negative level of welfare. The situation would not simply be one in which you are bereft of goods. When you are negative in welfare, some robust bads accrue to you. The mere lack of goods would result in a level of welfare that is, at worst, neutral. The presence of a robust bad is required to take your level of welfare to the negative. It is possible that the loss of a good is itself a robust bad and that such losses can cause your level of welfare to dip below the neutral point, depending on their magnitude and depending on how you were faring before you incurred the loss. But in those cases, such losses would count as robust bads rather than *mere* absences of the good.<sup>19</sup>

---

<sup>16</sup> Ibid. (p. 262).

<sup>17</sup> I thank David Sobel for suggesting this argument.

<sup>18</sup> Kagan 2014 (p. 263).

<sup>19</sup> Sumner (2020, pp. 420-421) and Rice (2019, p. 1074) acknowledge that robust bads—or what Sumner calls *substantive* bads—are required for negative welfare.



This observation can help us better characterize the concept of a robust bad. I suggest that we follow Kagan and say that robust bads are those states that directly reduce well-being but add that they are also capable of making a subject have a negative level of welfare. This new characterization handles nicely the case in which the mere absence of a pleasurable experience amounts to a reduction in well-being. *Mere* absences of pleasure—that is, absences of pleasure that are not also robustly bad for you—are capable of reducing well-being, but they are not capable of making a subject have a negative level of welfare, so mere absences of pleasure are correctly excluded from the class of robust bads on the favored proposal.

### III. The Dissonance-Constraint Argument

It is natural when in search of a rationale for subjectivism about ill-being to see whether the rationales that have been offered in support of subjectivism about well-being can be adapted for that purpose. In a seminal paper, Connie Rosati offers five arguments in favor of “internalism about a person’s good,” which is her preferred way of referring to subjectivism about well-being.<sup>20</sup> These arguments have been roundly criticized by others, so I will not discuss them here.<sup>21</sup> More recently, however, Dale Dorsey has offered an argument in favor of a doctrine that is closely related to subjectivism about well-being, and it pays to examine that argument to see if we can adapt it to support subjectivism about ill-being. I will argue that with the appropriate adjustments, Dorsey’s argument is suitable for that purpose.

Dorsey offers what he calls the *relationship-to-value argument* in favor of the following conclusion:

<b>Good-Value Link:</b>	For any object, event, state, etc., $\Phi$ and agent $x$ , $\Phi$ is good for $x$ only if and (at least in part) because $\Phi$ is valued, under conditions $c$ , by $x$ .
-------------------------	--

---

<sup>20</sup> See Rosati 1996 (pp. 309-324).

<sup>21</sup> See Sarch 2011 (section two) and Dorsey 2021 (pp. 91-96).

Note that, given certain plausible assumptions about the nature of valuing, the good-value link is not identical to subjectivism about well-being. The good-value link says that in order for a state of affairs to be good for you, you must value it under the proper conditions, but subjectivism about well-being only requires that you have some positive attitude or other, under the proper conditions, towards each state that is good for you. Subjectivism about well-being is thus a broader thesis about well-being, provided that we think of valuing as a special kind of positive attitude or as reducible to some subset of positive attitudes.<sup>22</sup> Not every subjective theory satisfies the good-value link since some positive attitudes are not plausibly construed as a valuing attitude.<sup>23</sup>

Dorsey's argument for the good-value link comes in two steps. The first step is to argue that there is what he calls a "kinship relation" between a welfare subject and the states that are basically good for that subject. I take it that this kinship relation is what I have been referring to throughout as resonance. A precise characterization of this resonance or kinship relation is elusive, but Dorsey give us examples, such as the positive relationship that obtains between "the spring budbreak and the *vitis vinifera*" and the positive relationship between "the speedy baserunner and the inside-the-park homerun."<sup>24</sup> Dorsey supports this first step by appealing to the very distinction between "prudential value and other forms of goodness." He writes:

---

<sup>22</sup> There are, of course, different theories of valuing. For Dorsey (2021, chapter five), for *S* to value *x* is for *S* to believe, under the proper conditions, that *x* is good for *S*. For Raibley (2013, p. 191), a person's valuing attitudes are "always partially constituted by some [adequately informed] diachronically stable and non instrumental pro-attitude" of theirs that they take "to be representative of who they are and who they want to be" and "that must inform or structure their emotional responses and practical deliberations." For Tiberius (2018, p. 40), valuing attitudes are "patterns of relatively robust desires and emotions that we endorse as giving us reasons relevant to planning and evaluating our lives."

<sup>23</sup> Some would insist, against the picture I have sketched here, that all subjectivists endorse the good-value link and that they just differ with respect to their preferred account of valuing. On this construal, "valuing" refers to that positive attitude or subset of attitudes, whatever they are, that speak for a subject or capture her point of view or represent, in some intended sense, *who she is*. But not all subjectivists think of themselves as giving a theory of valuing. Furthermore, I take it that it is an open question as to whether the attitude that fits this bill is the one whose satisfaction makes a subject better off. For example, a desire theorist might say that desire satisfaction is good for you but deny that your desires speak for you or capture your point of view or represent who you are. I do not think we would have to say that this person's views are incoherent. I thank Joseph Van Weelden and Teresa Bruno-Niño for discussion of this issue.

<sup>24</sup> Dorsey 2021 (p. 97).

Surely part of the difference here [...] is that for something to be of intrinsic value *for* a particular welfare subject, it should be the case that this thing bears a particular sort of relationship to that subject. After all, to say that  $\Phi$  is prudentially valuable for  $x$  isn't just to say that  $\Phi$  is *good*; it is also to say that  $\Phi$  is good *for*  $x$ ,  $\Phi$  is  $x$ 's *good*.<sup>25</sup>

Dorsey postulates that it is this positive connection, or resonance, that makes  $\Phi$   $x$ 's good as opposed to merely good *simpliciter*. What's more, resonance may not be univocal among types of subjects. What it is for a state to resonate with a nonhuman animal may very well differ from what it is for a state to resonate with a human being.

The next step for Dorsey is to argue that, for valuers in particular, “valuing attitudes *construct* the [resonance] relation.”<sup>26</sup> That is to say that when it comes to valuers, the subject valuing a state under the proper conditions is both necessary and sufficient to establish the resonance relation between that subject and that state.<sup>27</sup> Dorsey attempts to support this premise by pointing out that acquiring the capacity to value is a fundamental change in the subject. As he puts it, the subject “now has the capacity to [...] navigate the world of values for themselves.” Dorsey continues:

When an individual gains the capacity to value, this person now has the capacity to forge deeper *relationships* to particular values (or potential values) than would otherwise have been possible. [...] By exercising my valuing capacities (by *being a valuer*) I can mark out particular objects, states of affairs, and so forth as mattering to me, as important to me, as—in a very real way—*mine*, in a way that is not possible for non-valuing agents.<sup>28</sup>

Suppose that Dorsey is correct that becoming a valuer is a remarkable change for a welfare subject in precisely the way that he describes: the subject is now able to establish a kind of positive relationship with states of affairs in a way that non-valuing subjects cannot. But notice that this claim fails to establish Dorsey's premise. Even if it is true that a subject valuing a state of affairs is

---

<sup>25</sup> Ibid. (p. 96, emphasis in original).

<sup>26</sup> Ibid. (p. 98, emphasis in original).

<sup>27</sup> It should be pointed out that merely having the capacity to value is not sufficient to be a valuer in Dorsey's sense. According to Dorsey, a valuer actually exercises their capacity to value. A consequence of this claim is that the good-value link does not apply to beings who have the capacity to value but do not exercise that capacity, even if they are otherwise normal human adults. For discussion, see Dorsey 2021 (section 5.7).

<sup>28</sup> Dorsey 2021 (p. 98, emphasis in original).

sufficient to establish the resonance relation between that subject and that state, it surely does not follow that the subject doing so is *necessary* for the same.

Indeed, it is plausible that positive attitudes other than a valuing attitude can establish the resonance relation for valuers. Consider, for example, the following remarks from Peter Railton, which are widely regarded to be the starting point for theorizing about the resonance constraint:

It does seem to me to capture an important feature of the concept of intrinsic value to say that what is intrinsically valuable for a person must have a connection with what he would find in some degree compelling or attractive, at least if he were rational and aware. It would be an intolerably alienated conception of someone's good to imagine that it might fail *in any such way* to engage him.<sup>29</sup>

On a plausible interpretation, Railton implies that there are different ways for a person to find a thing compelling or attractive and that a person capable of doing so must find the states of basic benefit to that person compelling or attractive in at least some of these ways. Valuing a state of affairs is at least one way to find it compelling or attractive, but there are others, such as simply liking it or having a first-order desire for it. It is unlikely that the situation is different for the class of valuers than it is for the class of persons.

In light of these observations, I suggest that we modify the second step of Dorsey's argument to remain more ecumenical with respect to which positive attitudes are capable of establishing the resonance relation for beings like us. We might modify Dorsey's premises as follows:

- |                              |   |
|------------------------------|---|
| <b>Resonance Constraint:</b> | A token state of affairs, $x$ , is basically good for a welfare subject, $S$ , only if and (at least in part) because there is a positive connection, or a resonance, between $S$ and $x$ .   |
| <b>Positive Attitudes:</b>   | There is a resonance between a welfare subject who is capable of having positive attitudes and a state of affairs, $x$ , only if (and at least in part) because that subject has a positive attitude towards $x$ under the proper conditions. |

---

<sup>29</sup> Railton 1986 (p. 9, emphasis added).

These two premises support subjectivism about well-being more generally, rather than Dorsey's good-value link in particular:

**Subjectivism About Well-Being:** A token state of affairs,  $x$ , is basically good for a welfare subject who is capable of having positive attitudes only if and (at least in part) because  $x$  is suitably connected, under the proper conditions, to that subject's positive attitudes.

The basic reasoning of the argument is that for beings like us—beings capable of having positive and negative attitudes—our positive attitudes establish the kind of positive connection, or resonance, that must be present between any welfare subject and the token states that are of basic benefit to that subject. It follows that each state of basic benefit to beings like us must be connected in the right way to our positive attitudes. It is this modified version of Dorsey's relationship-to-value argument that we can adapt for our purposes of supporting subjectivism about ill-being.

Just as the resonance constraint can be used in conjunction with a plausible claim about a subject's positive attitudes in order to derive subjectivism about well-being, we can use what I call the *dissonance constraint* in conjunction with a plausible claim about a subject's negative attitudes to derive subjectivism about ill-being:

**Dissonance Constraint:** A token state of affairs,  $x$ , is basically bad for a welfare subject,  $S$ , only if and (at least in part) because there is a negative connection, or a dissonance, between  $S$  and  $x$ .

**Negative Attitudes:** There is a dissonance between a welfare subject who is capable of having negative attitudes and a state of affairs,  $x$ , only if (and at least in part) because that subject has a negative attitude towards  $x$  under the proper conditions.

Plausibly, at least part of what makes a state of affairs not just bad but bad *for you* is the negative relationship between you and that state. For beings like us with the capacity to forge negative

attitudes towards particular states of affairs, this negative relationship is established by our negative attitudes. This adaptation of Dorsey's argument is a plausible rationale for subjectivism about ill-being.

Furthermore, our intuitions about cases confirm that subjectivism about ill-being is true. Suppose that Stewart stubs his toe each morning on the corner of his bed frame. The first few times this happens, Stewart is quite bothered by and strongly averse to the sensation he gets on those occasions. He keeps stubbing his toe each morning to the point where he eventually gets used to the sensation he gets when he does so and has no negative attitude towards it. An observer who witnesses Stewart stub his toe might ask Stewart whether he is pained by, averse to, or otherwise bothered by the sensation he is currently feeling. Stewart would respond by saying that the sensation of stubbing his toe would previously cause him quite a bit of distress, but now he is used to it. "It's just another neutral sensation that is a regular part of my morning, not different in kind from the sensations I get when I brush my teeth or get dressed." It strikes me as overwhelmingly plausible that after he has become accustomed to the feeling, the sensation Stewart gets when he stubs his toe is not basically bad for him, and it seems to me that at least part of what explains this fact is that he does not have any negative attitudes towards it.

#### **IV. Bradford's Reverse Conditionalism**

Pain is perhaps the best candidate for a robust bad. When it comes to testing subjectivism about ill-being against our intuitions about what is of ultimate harm to beings like us, examining the badness of pain is a good place to start. Subjectivists about ill-being explain the badness of the experience of pain for a subject, at least in part, by appealing to that subject's negative attitudes. In a recent paper, Gwen Bradford has proposed a view about the badness of pain that is incompatible with this picture. In this section, I explain Bradford's view and her reasoning for it and argue that

she has given us insufficient reason to favor her view—a view that she calls *reverse conditionalism*—over a simpler, subjective theory of ill-being.<sup>30</sup>

Our question concerns the *value* of pain, not its *nature*.<sup>31</sup> Nonetheless, accounts of what pain *is* tend to mirror accounts of *why pain is bad for you*. When we investigate the nature of pain—that is, when we try to say what it is that makes an experience painful—we might appeal to the phenomenal character of the experience itself or to its relationship to the attitudes of the subject who experiences it. Similarly, when we try to give an account of the badness of pain, we can do so by appealing to the phenomenal character of the experience itself or to its relationship to the subject’s attitudes. These options give us four possible combinations of views: a proponent of a phenomenological theory of the nature of pain can adopt either a phenomenological or an attitudinal theory to explain its badness and a proponent of an attitudinal theory of the nature of pain can adopt either a phenomenological or an attitudinal theory to explain its badness. Note, however, that though these are four logically possible combinations of positions, not all of them are plausible. For example, attitudinal theories about the nature of pain allow for a phenomenologically diverse set of experiences to count as painful. Proponents of these theories presumably cannot avail themselves of a phenomenological theory of pain’s badness because doing so would ostensibly involve specifying a shared phenomenological character among painful experiences or at least a loose family resemblance among them.<sup>32</sup>

---

<sup>30</sup> I will follow Bradford and use the term ‘pain’ to refer to the broader class of unpleasant experiences (including experiences that some regard as unpleasant but not painful such as dizziness or nausea) and to encompass both physical and psychological pains.

<sup>31</sup> In what follows, I will assume that pain is bad and put the question of how to explain the value of pain in terms of explaining its badness.

<sup>32</sup> Bradford makes a similar point when she says that the proponent of what she calls a *desire-reductive account* of the nature of pain—which is a species of what I am here referring to as an attitudinal theory of pain—cannot explain the badness of pain by appealing to its phenomenal character. See Bradford 2020 (p. 237).

One example of a phenomenological theory of the badness of pain is what Bradford calls *dolorism*, the view that painful experiences are bad in virtue of their “negative feeling tone.”<sup>33</sup> Sometimes referred to as *hedonic tone*, feeling tone can be thought of as a feature of experience or a dimension along which experiences can vary.<sup>34</sup> An experience’s feeling tone can be either positive or negative, depending on whether the subject experiences it in a positive or negative way. Phenomenological theories of the badness of pain respect the commonsense view that pain is bad because of how it feels, and dolorism in particular respects the commonsense view that pain is bad because it hurts.

Dolorism is incompatible with subjectivism about ill-being since it does not explain the badness of pain in terms of the negative attitudes of the subject. The dolorist can adopt an attitudinal theory of pain, which, depending on the details of the theory, would guarantee that each token painful experience is the object of one of the subject’s negative attitudes. But this feature of the theory is not enough to make it compatible with subjectivism about ill-being. Our doctrine requires that the subject’s negative attitudes figure in the explanation of the badness, but according to dolorism paired with an attitudinal theory of the nature of pain, the subject’s negative attitudes are explanatorily idle when it comes to explaining the badness of pain. For the dolorist, irrespective of their favored theory of the nature of pain, the phenomenology of the experience does the heavy lifting when it comes to explaining the badness of pain.

Bradford rejects dolorism due to its inability to accommodate the existence of two phenomena. First, there are what Bradford calls *hurts-so-good* experiences, which are supposed to be experiences in which “we enjoy physical pain, such as the pain of exertion in intense athletic activity.”<sup>35</sup> Other examples of hurts-so-good experiences include “eating very spicy food, getting a

---

<sup>33</sup> Ibid. (p. 238).

<sup>34</sup> See Crisp 2006 (p. 109) and Kagan 1992 (p. 172).

<sup>35</sup> Bradford 2020 (p. 239).



deep tissue massage, jumping in a freezing lake, sitting in a very hot sauna, or eating something with a strong bitter or sour flavour.”<sup>36</sup> According to Bradford, such painful experiences are not bad for the people who experience them. Bradford also rejects dolorism on the basis that it implies, implausibly, that the pain of subjects with pain asymbolia—the medical condition in which one apparently experiences pain that is not unpleasant—is bad for them.<sup>37</sup>

A competitor to dolorism is conditionalism, an attitudinal theory of the badness of pain according to which pain’s badness is wholly explained by its relationship to the negative attitudes of the experiencing subject. Bradford prefers a sort of hybrid of these two options, which she calls “reverse conditionalism.” In this view, painful experiences are bad in virtue of how they feel *unless* the subject has a value-defeating positive attitude towards it. (A more apt name for the view might have been *dolorism with conditions*.) The view is taken to be an improvement on dolorism because it easily accommodates hurts-so-good experiences and pain asymbolia. According to reverse conditionalism, hurts-so-good experiences are those painful experiences towards which you have a positive attitude, so they are not bad for you. Moreover, those with pain asymbolia have a positive attitude of *unbotheredness* that undermines what otherwise would have been the badness of their painful experiences.<sup>38</sup>

Note that conditionalism is a subjective theory of ill-being because it says that the subject’s negative attitude towards the painful experience is both necessary and sufficient for its badness.<sup>39</sup> Reverse conditionalism, however, is incompatible with subjectivism about ill-being. To see why,

---

<sup>36</sup> Ibid. (p. 239).

<sup>37</sup> Ibid. (p. 242).

<sup>38</sup> It is unclear whether unbotheredness is properly understood as an attitude at all, and even if it is to be understood as an attitude, it is unclear that it should be understood as a *positive* attitude that would have the power, as alleged by Bradford, like that of other characteristically positive attitudes such as desire to undermine the badness of painful experiences.

<sup>39</sup> Strictly speaking, conditionalism, as Bradford gives it, is just a view about the badness of *experiences*, so whether conditionalism is a subjective theory of ill-being will depend on whether the conditionalist says that anything other than experiences can be basically bad for you and whether the theory implies that they can be bad for you irrespective of your negative attitudes. I elide this detail in the main text for the sake of simplicity.

suppose that you have a painful experience with a negative feeling tone and that you do not have a positive attitude towards the experience. Reverse conditionalism implies that the painful experience is bad for you and that its badness is explained by how it feels and not at all by its connection to your negative attitudes.

Why does Bradford prefer reverse conditionalism, a view that is incompatible with subjectivism about ill-being, over conditionalism, a simpler, subjective theory of ill-being? Bradford expresses three concerns about conditionalism. First, she argues that “there are experiences that conditionalism assesses as bad which we do not intuitively think are bad.”<sup>40</sup> Bradford’s example is a toddler who dislikes the experience of his mother’s kiss. Bradford’s intuition is that this experience is *not* bad for the toddler due to its positive feeling tone. But it is not at all clear that the experience of his mother’s kiss would have a positive, rather than a negative, feeling tone, provided that the toddler dislikes it. Plausibly, the fact that an experience is disliked can at least sometimes make its feeling tone negative. The feeling tone of a subject’s experience does not float entirely free of that subject’s attitudes towards it.<sup>41</sup> The conditionalist cannot appeal to this negative feeling tone when explaining the badness of the mother’s kiss since the conditionalist can only offer such explanations that appeal to the negative attitudes of the subject, but if the experience of the mother’s kiss has a negative feeling tone due to the toddler’s negative attitude towards it, then Bradford’s case for thinking the mother’s kiss is not bad for the toddler is thereby undermined.

Moreover, even if the experience of his mother’s kiss has a positive feeling tone, it is plausible that it is bad for the toddler at least in part because of the toddler’s negative attitude

---

<sup>40</sup> Bradford 2020 (p. 245).

<sup>41</sup> Bradford (2020, p. 247) appears to acknowledge as much when she writes that “attitudes can mitigate the degree to which an experience is felt as unpleasant - an unexpected pinprick feels more unpleasant and alarming than a pinprick that is expected and contextualized.” Also, note that I am not claiming that the feeling tone of a subject’s experience is determined entirely by their attitudes towards it, nor am I claiming that a negative attitude towards an experience always makes the experience have a negative feeling tone. I only claim that in some cases having a negative attitude towards an experience can have this effect.

towards it. There is little rational pressure to accept the view that experiences with a positive feeling tone can never be made bad for us at least in part due to the negative attitudes we take towards them.<sup>42</sup> After initially stating the objection, Bradford herself sides with the conditionalist response to this first objection.<sup>43</sup>

Bradford's second concern about conditionalism is that it "cannot capture [the badness of] displeasure for psychologically unsophisticated welfare subjects."<sup>44</sup> She writes:

Perhaps a toddler is simply not capable of forming the relevant higher-order attitudes for establishing the goodness or badness of an experience. The pain of non-human animals [such as toads or shrimps] is bad for them, but it is unclear that all non-human animals have higher-order attitudes. Further examples include extremely intense physical pain or extremely intense unpleasant psychological experience that limit one's cognitive abilities to the point of eclipsing any ability to form attitudes. Such mind-blowing pain leaves no mental space to take an attitude toward it, one might think.<sup>45</sup>

We might tease apart two distinct objections here. The first is that conditionalism is unable to accommodate the fact that pain is bad for cognitively unsophisticated subjects like toddlers and certain non-human animals.<sup>46</sup> The second is that the theory cannot accommodate the badness of mind-blowing pain towards which one is allegedly precluded from having a negative attitude.

---

<sup>42</sup> The case of the mother's kiss undermines the conditionalist's claim that having a negative attitude towards an experience is sufficient to make it bad for you, but the subjectivist about ill-being need not accept this claim. My aim in the main text is to undermine Bradford's overall case against conditionalism to show that she has not supplied good reasons to prefer reverse conditionalism over conditionalism, and doing so involves, incidentally, defending conditionalism against Bradford's objections that would not also apply to subjectivism about ill-being as such. It is thus important to note that even if Bradford's first concern about conditionalism were well-founded, it would only undermine conditionalism *per se*, not subjective theories of ill-being as a class.

<sup>43</sup> Bradford 2020 (p. 245).

<sup>44</sup> *Ibid.* (p. 248).

<sup>45</sup> *Ibid.* (p. 245). There is an additional element of Bradford's objection that I do not discuss in the main text. She says that whether toddlers and certain non-human animals have sophisticated attitudes is an empirical question and that it is preferable not to rest the truth of conditionalism on this empirical question. The best formulation of the objection, however, does not rely on this empirical question. As long as there are possible worlds in which such beings do not have sophisticated attitudes, then the theory can be accused of failing to deliver the intuitively correct verdicts about their welfare in those worlds.

<sup>46</sup> As Bradford notes, Lin (2017b) gives a similar objection, which I discuss in section five below. My response to each is the same: the subjectivist about ill-being can and should say that the cognitively unsophisticated attitudes of the cognitively unsophisticated subject can explain why some token state is bad for them.

To address the first objection, the conditionalist can appeal to cognitively unsophisticated negative attitudes in order to explain the badness of painful experiences of cognitively unsophisticated subjects. Normal human adults are capable of having cognitively sophisticated negative attitudes such as disvaluing, whereas newborn babies may be capable of only forming relatively primitive negative attitudes such as aversions for experiences they dislike. The conditionalist can say that the painful experiences of cognitively unsophisticated subjects is explained by the cognitively unsophisticated negative attitudes they take towards those experiences. It may very well be that the badness of the painful experiences of cognitively sophisticated subjects is explained by their cognitively sophisticated negative attitudes while the painful experiences of cognitively unsophisticated subjects is explained differently due to their different nature.<sup>47</sup>

The conditionalist can address the second objection by denying that mind-blowing pain leaves no mental space for the subject who experiences it to take the relevant kind of negative attitude towards it. Perhaps it is true that the mind-blowing pain prevents the development of a cognitively sophisticated negative attitude towards it, but it is less than clear that it prevents the development of a cognitively unsophisticated attitude. Furthermore, the objection relies on a claim concerning the temporal relationship between the painful experience and the negative attitudes that the conditionalist need not accept. It is natural to interpret conditionalism as requiring that the subject have the relevant negative attitude at the time at which the painful experience occurs. This version of the view is subject to the present objection because the mind-blowing pain had at time  $t$  is thought to prevent the experiencing subject from having the negative attitude at time  $t$  that would render the painful experience bad for the subject. But the conditionalist can instead formulate a different version of the view according to which there need not be temporal overlap between painful experiences and the negative attitudes that make them bad for the experiencing subject. On one way

---

<sup>47</sup> I argue in section five below, however, that the subjectivist about ill-being should formulate their view only in terms of unsophisticated negative attitudes, even as it applies to sophisticated welfare subjects.

of specifying the details, the conditionalist can say that the badness of a painful experience can be explained by the negative attitudes that the subject had at some time earlier than  $t$ . This view allows the conditionalist to say that the mind-blowing pain is bad for the subject at  $t$  even though the subject does not have any negative attitudes towards it at  $t$ .

Bradford's third concern with conditionalism is that "it misses something important: the quality of the feeling of unpleasantness has no role in the explanation of its badness."<sup>48</sup> Pain is bad for you because *it hurts*, but conditionalism cannot accommodate this fact because the view does not allow for the phenomenal character of pain to explain its badness. The objection might strike one as somewhat question-begging since it does not amount to much more than the claim that conditionalism should be rejected because it is not a form of dolorism. More charitably, we can interpret the concern as saying that conditionalists are unable to accommodate a commonsense datum about the badness of pain, namely that the phenomenal character of painful experiences must at least partially explain the badness of those experiences.<sup>49</sup>

The conditionalist can respond to this objection by distinguishing between two different ways that the phenomenal character of a painful experience can explain the badness of that experience: either *directly* or *indirectly*. The dolorist says that the phenomenal character of a painful experience directly explains the badness of that experience. The phenomenology of painful experiences cannot play such a direct role on the conditionalist's account, but it can play an indirect explanatory role by explaining something else—namely, the relevant negative attitude—which itself directly explains the badness of the experience.

---

<sup>48</sup> Bradford 2020 (p. 245).

<sup>49</sup> As with Bradford's first concern, this third concern does not touch subjectivism about ill-being as such. Unlike conditionalism, other subjective theories of ill-being can say that the phenomenal character can, in the way that I go on to explain in the main text, directly explain the badness of the experience (together, of course, with the fact that the subject has a negative attitude towards it).

To see why, consider the fact that the conditionalist can say that painful experiences are bad for you if and only if you have an intrinsic, *de re* negative attitude towards them. In fact, the conditionalist has independent reason to formulate their view in this way in order to avoid certain counterexamples.<sup>50</sup> The negative attitudes that explain the badness of the painful experience according to the conditionalist must be negative attitudes that are directed towards the experience for its own sake and in virtue of how it feels. In other words, for the conditionalist, as Bradford herself puts it, “the *res* of the *de re* is the phenomenal character of the experience” itself.<sup>51</sup> Bradford might insist that the commonsense datum about the badness of pain on which the objection relies is that the phenomenology of the painful experience directly, rather than merely indirectly, explains the badness of the experience, but we have been provided no reason for thinking this is true. Furthermore, it is doubtful that common sense distinguishes between these two possible explanatory roles, even less likely that it has a preference for one over the other, and less likely still that the explanatory role that common sense prefers is the one that happens to favor dolorism over conditionalism.

In this section, I have shown that there is little rational pressure to accept Bradford’s reverse conditionalism over a more standard form of conditionalism according to which the badness of a painful experience is conditional on one’s having a negative attitude towards it. Though Bradford’s reverse conditionalism is incompatible with subjectivism about ill-being, we have no reason to prefer it over conditionalism, which is a simpler, subjective theory of ill-being.

---

<sup>50</sup> For example, you might have a negative attitude towards the pleasant smell of freshly baked cookies only because you do not want to be tempted to eat the cookies. See Bradford 2020 (p. 243). Heathwood (2007, pp. 28-32) provides other examples to motivate these same refinements to a desire theory of the nature of sensory pleasure.

<sup>51</sup> Bradford 2020 (p. 243).

## V. Lin's Arguments Against Subjectivism

In an influential paper, Eden Lin offers a collection of arguments against different varieties of subjective theories of welfare.<sup>52</sup> Recall that a subjective theory of welfare implies both subjectivism about well-being and subjectivism about ill-being. Some of Lin's arguments are formulated as arguments against subjective theories of well-being but it is not difficult to formulate versions of them that target subjective theories of ill-being as well. In this section, I will defend unsophisticated subjective theories of ill-being—theories that appeal only to the unsophisticated negative attitudes of the subject, such as the subject's aversions—from Lin-style arguments.

Lin's taxonomy of subjective theories of welfare is instructive.<sup>53</sup> Subjective theories imply that a state of affairs is good (bad) for you only if it is suitably connected, under the proper conditions, to your positive (negative) attitudes. As I noted earlier, proponents of such theories will disagree over which attitudes are relevant and how to specify the proper-conditions clause. With respect to disagreements over the relevant attitudes, the main options are sophisticated attitudes like evaluative beliefs and (dis)valuing or unsophisticated attitudes like desires (aversions). With respect to disagreements over the proper conditions, the main options are to specify them as the conditions of the world in which the subject exists (this is what Lin calls *same-world* subjectivism) or the conditions of a world in which the subject is idealized in some respect (this is what Lin calls *ideal-world* subjectivism). If we focus solely on the sophisticated attitudes of evaluative beliefs and (dis)valuing and the unsophisticated attitude of desire (aversion), this taxonomy generates six subjective theories of welfare: both a same-world and an ideal-world version of each a judgment-based, a value-based,

---

<sup>52</sup> Lin 2017b.

<sup>53</sup> Ibid. (pp. 356-357).

and a desire/aversion-based subjectivism. I will bracket ideal-world subjective theories due to recent arguments that the subjectivist need not and should not idealize.<sup>54</sup>

Lin's argument against subjective theories of welfare begins with an argument against sophisticated theories. Take your favored sophisticated theory. For the purposes of illustration, we can use a judgment-based subjectivism about welfare:

<b>Same-World Judgment Subjectivism About Welfare:</b>	$x$ is basically good (bad) for you at $W$ if and only if, because, and to the extent that at $W$ , you believe that $x$ is basically good (bad) for you.
--	---

Either the proponent of this view thinks the theory is true of cognitively unsophisticated subjects like newborn babies or it says that a different theory is true of these subjects. If it is claimed that the theory is true of them, then we can easily construct counterexamples in which it seems clear that a newborn baby is benefited (harmed) despite the fact that the baby does not believe that the beneficial state is good (bad) for themselves due to their inability to have such attitudes. Suppose instead that it is claimed that the theory is true of normal human adults but not of newborn babies. Then, as Lin puts it, "adult welfare diverges from neonatal welfare in a way that cannot plausibly be explained."<sup>55</sup>

Indeed, I have argued elsewhere that other cases present an even stronger case against the theory.<sup>56</sup> The judgment subjectivist can attempt to explain this divergence between adult welfare and neonatal welfare by appealing to the fact that becoming a being capable of forming welfare beliefs is a change of tremendous significance. Notice, however, that they could not make such an appeal to

---

<sup>54</sup> For criticisms of idealization, see Rosati 1995 and Enoch 2005. For an argument that subjectivists need not idealize, see Lin 2019. Despite these criticisms, it should be noted that whether the subjectivist should idealize is still a live debate.

<sup>55</sup> Lin 2017b (p. 358). Notice that Lin's argument reinforces my earlier point that Dorsey has good reason to reformulate his relationship-to-value argument as an argument in favor of subjectivism about well-being more generally rather than as an argument for the good-value link in particular. Valuing is a sophisticated attitude, and sophisticated subjective theories of welfare are the target of Lin's argument. I argue below that subjectivists about welfare can blunt the force of Lin's arguments if they formulate their view solely in terms of unsophisticated attitudes.

<sup>56</sup> See Kelley 2021 (pp. 301-303).



explain how the welfare of normal human adults diverges from the welfare of otherwise normal human adults who do not have welfare beliefs since such a divergence wears its implausibility on its face. Similar reasoning undermines other sophisticated theories. Subjectivists should therefore formulate their view in terms of unsophisticated attitudes.<sup>57</sup>

Lin's argument against sophisticated subjective theories of welfare seems to equally undermine both the subjective theories of well-being and ill-being that compose such theories. Just as newborn babies are capable of being benefited despite lacking sophisticated attitudes, they are also capable of being harmed. Furthermore, a divergence between normal human adult well-being and neonatal well-being such as the one that Lin identifies would call out for explanation just as loudly as an analogous divergence between normal human adult ill-being and neonatal ill-being.

However, as I will demonstrate, Lin's argument against unsophisticated theories of welfare does not have this feature. In what follows, I will explain Lin's argument against unsophisticated subjective theories of well-being and show that an analogue of his argument does not undermine subjective theories of ill-being. Furthermore, if we assume a certain symmetry between well-being and ill-being, my argument also shows that Lin's argument does not undermine subjective theories of well-being either. Though I will have achieved my aims if I show that Lin's argument does not undermine subjective theories of ill-being in particular, the upshot of these considerations is that

---

<sup>57</sup> Yelle (2016) defends sophisticated theories of welfare against Lin's objection on the grounds that persons have different dimensions of welfare. According to Yelle, "a subject can have multiple levels or dimensions of welfare which correspond to the different kinds to which she belongs, e.g., 'human being', 'person', 'experiencing subject', etc." (p. 1410). This suggestion allows the sophisticated subjectivist to say that there is a sense in which there is no divergence between neonatal and normal human adult welfare provided that we are focused on normal adult human welfare *qua* human. The sense in which there is a divergence is as follows: unlike normal human adults, newborn babies do not have a welfare *qua* persons since they lack the requisite capacities. But this divergence, Yelle argues, is plausibly explained by the new and more sophisticated kind to which the subject belongs once they have reached the relevant point in their development. This response, however, does not take into account the fact that even among the class of persons, there are some who do not exercise the capacities in virtue of which they are persons. Either sophisticated subjective theories yield implausible verdicts about the welfare of such individuals or they imply that there is an implausible divergence between their welfare, *qua* person, and the welfare, *qua* person, of normal human adults who exercise the relevant capacities. So even if Yelle's proposal is effective as a response to Lin's argument, it has no purchase against this strengthened version of the argument.

Lin's argument does not undermine subjective theories of well-being either, thus leaving subjective theories of welfare in general entirely unscathed.

Consider an unsophisticated subjective theory of welfare such as the desire theory. It is the conjunction of the following two claims:

**Same-World Desire  
Subjectivism About Well-Being:**

$x$  is basically good for you at  $W$  if and only if, because, and to the extent that at  $W$ , you desire  $x$ .

**Same-World Aversion  
Subjectivism About Ill-Being:**

$x$  is basically bad for you at  $W$  if and only if, because, and to the extent that at  $W$ , you have an aversion to  $x$ .

Note that Lin's argument against sophisticated theories like judgment subjectivism does not undermine an unsophisticated theory like the desire theory since it is plausible that newborn babies are capable of unsophisticated attitudes like desires and aversions.

Lin's argument against the desire theory begins with the following observation. Cognitively unsophisticated subjects like newborn babies are capable of having only *experiential* desires, which are desires for an experience formed solely on the basis of the phenomenal character of that experience. It is plausible that a newborn can desire the experience of his father's warm embrace, but it is not plausible that they can desire that their experience of being embraced by their father be caused by their father actually embracing them rather than some other cause since such an attitude would require an understanding of causation beyond the grasp of a cognitively unsophisticated subject.<sup>58</sup>

Lin asks us to imagine a pair of newborn babies, Adam and Bill, who have the same desires and who are phenomenological duplicates. Both Adam and Bill have experiences that each desires, such as the experience of their father's warm embrace, but Adam is, for example, actually being embraced whereas Bill is hooked up to an experience machine and being fed a mere simulation of

---

<sup>58</sup> As Lin points out, the argument does not rely on the empirical claim that newborns have only experiential desire. Even if it turns out that newborns in the actual world have non-experiential desires, if there are newborns in merely possible worlds who have only experiential desires, the theory would deliver the (allegedly) intuitively incorrect judgments about their welfare. See Lin 2017b (p. 364).

being embraced. Keeping all else equal and focusing only on the neonatal period before either Adam or Bill is capable of forming desires for states beyond their own experiences, Lin judges that “Adam’s total welfare during the neonatal period is at least somewhat higher than Bill’s.”<sup>59</sup> As Lin puts it, “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.”<sup>60</sup>

Lin says that the desire theory is implausible because it implies that Adam and Bill are faring equally well. Lin writes:

[I]f 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. [...] [The fact that Adam’s total well-being during the neonatal period is at least somewhat higher than Bill’s] cannot be explained by a difference in how well-satisfied Adam’s and Bill’s desires are. But according to Same-World Desire Subjectivism, only such a difference could account for a difference in their welfare.<sup>61</sup>

Because the desire theory says that you have to desire a state in order for it to be good for you and since Adam and Bill have the same desires, the alleged difference in their total welfare during the neonatal period cannot be attributed to a difference in desire satisfaction. Though the example uses the desire theory and the unsophisticated attitude of desire, the argument could be run against a different subjective theory that appealed to some other unsophisticated attitude.

Notice that Lin’s argument only explicitly targets the claim that desire satisfaction is the sole basic good. Recall that Lin says the desire theorist cannot explain the alleged difference between Adam’s and Bill’s welfare by appealing to a difference in desire satisfaction. This remark suggests that he conceives of the case as one in which there is some basically good particular that accrues to Adam that does not accrue to Bill, which, if true, is a fact that the desire theorist cannot explain

---

<sup>59</sup> Ibid. (p. 363).

<sup>60</sup> Ibid. (p. 362-363).

<sup>61</sup> Ibid. (p. 363-364).

since it is a feature of the case that the two subjects have the same desires.<sup>62</sup> In other words, the argument only explicitly attempts to undermine the subjective theory of well-being that underwrites the desire theory and other unsophisticated theories of welfare.

We can nonetheless construct a case analogous to the case of Adam and Bill to leverage a Lin-style argument against subjective theories of ill-being. Imagine two newborns, Adam\* and Bill\*, both of whom have the same desires and aversions and are phenomenological duplicates. Just as it is plausible that newborns only have experiential desires, it is plausible that they only have experiential aversions, which are aversions for an experience formed solely on the basis of the phenomenal character of that experience. Suppose that Adam\* and Bill\* both have the experience of being tortured and that each is averse to this experience. Suppose further that Adam\* is actually being tortured but that Bill\* is merely having the simulated experience of being tortured.<sup>63</sup> The Lin-style argument against subjective theories of ill-being runs as follows. Intuitively, keeping all else equal and focusing only on the neonatal period before either is capable of forming aversions for states beyond their own experiences, Adam\* is faring worse than Bill\* during the neonatal period because there is some basically bad particular that accrues to Adam\* that does not accrue to Bill\*. The subjective theory of ill-being that underwrites the desire theory—the claim that aversion satisfaction is the only thing that is basically bad for you—cannot accommodate this fact since there is no difference in terms of aversion satisfaction between the two subjects. And since our selection of this particular theory of ill-being that appeals to the particular unsophisticated attitude of aversion was entirely

---

<sup>62</sup> He does say, however, that Bill is “*pitiful* in a way that Adam is not,” which suggests that perhaps he has in mind that there are some basically bad particulars that accrue to Bill that do not accrue to Adam. See Lin 2017b (p. 363, emphasis in original).

<sup>63</sup> I assume that all torturers are malicious agents. One might be worried that being given the experience of being tortured *just is* to be tortured. But I take it that to torture someone is not just to cause them to experience agonizing pain. Instead, it is to do so for a specific set of reasons (e.g., because you want them to suffer or because you want to extract some information from them). I am assuming that no one is simulating Bill\*’s experience of torture for any of these reasons, so Bill\* is not actually being tortured. I thank Eden Lin for discussion of this point.

arbitrary, the same kind of reasoning seems to apply with equal force to other unsophisticated theories of ill-being that appeal to other unsophisticated negative attitudes.

It is very intuitively compelling, however, that Adam\* is *not* faring worse than Bill\* during the neonatal period. Surely having the experience of being tortured but not actually being tortured is no better for you than having the experience of being tortured and actually being tortured.<sup>64</sup> It would be cold comfort to learn that you were merely having the experience of being tortured but not actually being tortured. If I were deciding whether to give a newborn Adam\*'s or Bill\*'s life and if I were taking into account only considerations of the newborn's welfare during the neonatal period, I would simply flip a coin to decide between the two. These considerations suggest that there are no basically bad particulars that accrue to Adam\* that do not accrue to Bill\*. Thus, even if Lin's argument undermines unsophisticated subjective theories of well-being, a similar style of reasoning does not undermine unsophisticated subjective theories of ill-being.

Our aim is to defend subjective theories of ill-being, and the reasoning just provided is sufficient for that purpose. But it is worth pointing out that the fact that the Lin-style reasoning does not undermine unsophisticated subjective theories of ill-being strongly suggests that Lin's argument does not undermine unsophisticated subjective theories of well-being either. The truth of this claim is due to the intuitive symmetry between well-being and ill-being. If there are no basically bad particulars that accrue to Adam\* that do not accrue to Bill\*, then there are no basically good particulars that accrue to Adam that do not accrue to Bill. In other words, if a case were insufficient to show that there are basic bad kinds in addition to aversion satisfaction, it would be surprising if an analogous case were sufficient to show that there are basic good kinds in addition to desire satisfaction. Our allegiance to this symmetry need not be absolute. Further argumentation may very well shake our confidence, but in the absence of any reason to reject it, it is reasonable to assume

---

<sup>64</sup> Both Sumner (2020, p. 424) and Feldman (2004, p. 111) express similar intuitions about similar cases.

that our theories of welfare will respect this symmetry. Our focus is on defending subjective theories of ill-being in particular, but it is an appealing feature of my argument that it can get us a defense of subjective theories of welfare virtually for free.

## VI. Conclusion

In this paper, I have extended the popular, yet controversial subjectivist tradition concerning well-being to cover the new and burgeoning field of ill-being. I have done so by arguing for subjectivism about ill-being, which is the view that in order for a token state of affairs to be basically bad for you, it must be suitably connected, under the proper conditions, to your negative attitudes. I have shown that subjectivism about ill-being can be derived from a more general doctrine requiring a negative connection, or dissonance, between any welfare subject and the token states of affairs that are basically bad for that subject. I have defended a particular subjective theory of ill-being from the objections raised against it by Gwen Bradford, and I have defended unsophisticated subjective theories of ill-being from an objection inspired by Eden Lin.

I conclude the paper by noting one important way that we might incorporate subjectivism about ill-being into our theorizing about welfare. Objectivists about well-being who wish to adopt subjectivism about ill-being must reconsider their theory of well-being. After all, if subjectivism about ill-being is true, subjectivism about well-being is likely to be true as well. Moreover, subjectivists about well-being should take special care when attempting to extend their theory to cover the case of ill-being. For example, standard formulations of the desire theory say that getting what you want is good for you and *not* getting what you want is bad for you. But as I have argued elsewhere, this is a mistake.<sup>65</sup> This formulation of the theory violates subjectivism about ill-being since the theory implies that  $\sim p$  is basically bad for you, provided that you desire that  $p$ , irrespective of whether you have a negative attitude towards  $\sim p$ . Philosophers working on welfare should take

---

<sup>65</sup> See Kelley 2020 (chapter four).

care to ensure that each thing their theory says is basically bad for a subject is a thing towards which the subject is guaranteed to have a negative attitude.

## Bibliography

- Bradford, Gwen. "The Badness of Pain." *Utilitas*. Vol. 32, No. 2 (June 2020): 236-252.
- . "Perfectionist Bads." *The Philosophical Quarterly*. Vol. 71, No. 3 (July 2021): 586-604.
- Brandt, Richard B. *A Theory of the Good and the Right*. Oxford: Clarendon, 1979.
- Crisp, Roger. *Reasons and the Good*. Oxford: Oxford University Press, 2006.
- Dorsey, Dale. *A Theory of Prudence*. Oxford: Oxford University Press, 2021.
- Enoch, David. "Why Idealize?" *Ethics*. Vol. 115, No. 4 (July 2005): 759-787.
- Feldman, Fred. *Pleasure and the Good Life*. Oxford: Oxford University Press, 2004.
- Fletcher, Guy. "A Fresh Start for the Objective-List Theory of Well-Being." *Utilitas*. Vol. 25, No. 2 (June 2013): 206-220.
- Heathwood, Chris. "The Reduction of Sensory Pleasure to Desire." *Philosophical Studies*. Vol. 133, No. 1 (March 2007): 23-44.
- . "Desire-Fulfillment Theory." *The Routledge Handbook of Philosophy of Well-Being*, ed. G. Fletcher. New York: Routledge, 2016. 135-147.
- Kagan, Shelly. "Well-Being as Enjoying the Good." *Philosophical Perspectives*. Vol. 23, No. 1 (Dec. 2009): 253-272.
- . "An Introduction to Ill-Being." *Oxford Studies in Normative Ethics*, Vol. 4, ed. M. Timmons. Oxford: Oxford University Press, 2014. 261-288.
- . "What is the Opposite of Well-Being?" In *Principles and Persons: The Legacy of Derek Parfit*, eds. J. McMahan, et. al. Oxford: Oxford University Press, 2021.
- Kelley, Anthony. *Well-Being and Alienation*. 2020. Dissertation submitted to the University of Colorado, Boulder.
- . "The Welfare-Nihilist Arguments Against Judgment Subjectivism." *Journal of Ethics & Social Philosophy*. Vol. 19, No. 3 (March 2021): 291-310.



- Lin, Eden. "Enumeration and Explanation in Theories of Welfare." *Analysis*. Vol. 77, No. 1 (Jan. 2017a): 65-73.
- . "Against Welfare Subjectivism." *Nous*. Vol. 51, No. 2 (June 2017b): 354-377.
- . "Why Subjectivists About Welfare Needn't Idealize." *Pacific Philosophical Quarterly*. Vol. 100, No. 1 (March 2019): 2-23.
- Mathison, Eric. *Asymmetries and Ill-Being*. 2018. Dissertation submitted to the University of Toronto.
- Parfit, Derek. *Reasons and Persons*. Oxford: Oxford University Press, 1984.
- Raibley, Jason. "Values, Agency, and Welfare." *Philosophical Topics*. Vol. 41, No. 1 (Spring 2013): 187-214.
- Railton, Peter. "Facts and Values." *Philosophical Topics*. Vol. 14, No. 2 (Fall 1986): 5-31.
- Rawls, John. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.
- Rice, Christopher M. "Objective List Theories and Ill-Being." *Ethical Theory and Moral Practice*. Vol. 22, No. 5 (Nov. 2019): 1073-1085.
- Rosati, Connie. "Persons, Perspectives, and Full Information Accounts of the Good." Vol. 105, No. 2 (Jan. 1995): 296-325.
- . "Internalism and the Good for a Person." *Ethics*. Vol. 106, No. 2 (Jan. 1996): 297-326.
- Sarch, Alexander. "Internalism about a Person's Good: Don't Believe It." *Philosophical Studies*. Vol. 154, No. 2 (June 2011): 161-184.
- Sidgwick, Henry. *The Methods of Ethics*. (1874) London: Macmillan & Co., 1907.
- Singer, Peter. *Practical Ethics*. Cambridge: Cambridge University Press, 1979.
- Sumner, Wayne. "The Worst Things in Life." *Grazer Philosophische Studien*. Vol. 97, No. 3 (Aug. 2020): 419-432.
- Tiberius, Valerie. *Well-Being as Value Fulfillment*. Oxford: Oxford University Press, 2018.
- Tully, Ian. "Depression and the Problem of Absent Desires." *Journal of Ethics & Social Philosophy*. Vol.

11, No. 2 (Feb. 2017): 1-15.

Van Weelden, Joseph. "On Two Interpretations of the Desire-Satisfaction Theory of Prudential Value." *Utilitas*. Vol. 31, No. 2 (June 2019): 137-156.

Yelle, Benjamin. "In Defense of Sophisticated Theories of Welfare." *Philosophia*. Vol. 44, No. 4 (Dec. 2016): 1409-1418.