

# Welfare Invariabilism

We are not the only entities that are capable of well-being. Besides cognitively typical human adults, many other beings can fare well or poorly: seriously cognitively disabled adults, children, infants, non-human animals of various kinds, and perhaps even some plants. Philosophers usually do not specify which subjects their theories of welfare are supposed to cover. But it is natural to interpret them as theorizing, at least primarily, about beings like us: cognitively typical human adults.

*Invariabilism* is the view that the same theory of welfare is true of every welfare subject. *Variabilism* is the view that invariabilism is false. Because extant discussions have focused on typical human adults, and because the same theory is presumably true of all of us, little attention has been paid to whether variabilism or invariabilism is true. If variabilism is true, then the task of discovering the correct theory of a given subject’s welfare is a relatively low-stakes affair: to say that hedonism is true of your dog is not to deny that desire satisfactionism or perfectionism is true of you. If invariabilism is true, however, then we cannot proceed in such a piecemeal fashion: to settle on a theory of one subject’s welfare is to accept that theory for all subjects.

In light of how many welfare subjects there are and how greatly they differ in their natures and capacities, it is natural to suppose that variabilism is true. I will argue that these considerations do not support variabilism, and indeed, that we should accept invariabilism. As I will explain, this has important implications: it eliminates many of the going theories of welfare while making some of the remaining ones more attractive.

## 1. Theories of Well-Being

First, let me clarify what is at issue in the debate between variabilism and invariabilism by explaining what a theory of well-being involves.

It is analytic that a subject’s level of welfare is determined by how good or bad for it everything is: the more particular things are good for a subject and the better for it they are, the higher its welfare, whereas the more particular things are bad for a subject and the worse for it they are, the lower its

welfare. But it would be double-counting to count things that are *merely derivatively* good or bad for a subject—good or bad for it solely in virtue of being appropriately related to other things that are good or bad for it (e.g., as causes or composites of those things). A subject’s level of welfare is fixed by how non-derivatively or *basically* good or bad for it everything is. Thus, any theory of welfare must perform two functions: it must enable us to identify the particular things that are basically good or bad for the relevant subjects, and it must tell us to how calculate the *basic prudential values* of those things (i.e., *how* basically good or bad for the subject they are).<sup>1</sup>

There is another function that any theory of welfare must perform: it must group the basically good or bad particulars into basically good or bad *kinds*—or, as I call them, *basic goods and bads*. Hedonism, for instance, claims that pleasure is the only basic good, whereas an objective list theory might say that pleasure, friendship, and knowledge are all basic goods. This function of a theory of welfare is an extension of the first one: a theory enables us to identify the basically good or bad particulars precisely by identifying the basically good or bad kinds. According to desire satisfactionism, for example, what enables us to identify the particulars that are basically good for you is the fact that the kind *thing that satisfies one of the subject’s desires* is the only basic good. Why is it basically good for you to be listening to jazz even though it isn’t basically good for you to be chatting with your mother-in-law? Because you desire the former but not the latter, and things that satisfy one of your desires are the only basic good.<sup>2</sup>

Finally, as this last example illustrates, every theory of welfare will invoke the basic goods and bads that it identifies to *explain* why the particulars that are basically good or bad have that status.<sup>3</sup> Desire satisfactionists claim that every particular that is basically good for you has that status because you desire it. Hedonists claim that every basically good particular has that status because it is a pleasure. An objective list theorist who says that pleasure and friendship are both basic goods will claim that some particulars are basically good because they are pleasures and others are basically good because they are instances of friendship. Thus, each basic good will correspond to a *good-making property*: a property that grounds or explains the basic goodness of the particulars that possess it. This is why

---

<sup>1</sup> I will proceed as though basic prudential values are numbers, but this is an idealization: in reality, many basically good or bad particulars may be only roughly comparable in value.

<sup>2</sup> Some desire satisfactionists prefer to say that *combinations* consisting of things that you desire and your desiring them (i.e., facts of the form *p & you desire p*) are the only basic good: see, e.g., Heathwood (2005). My remarks could easily be adapted to cover such “combo” desire satisfactionist views.

<sup>3</sup> In Lin (forthcoming), I clarify the sense in which theories of welfare are explanatory.

desire satisfactionism is a monistic theory (i.e., one on which there is exactly one basic good and one basic bad) even though many different kinds of thing are desired. Because desire satisfactionism says that every particular that is basically good for you has that status in virtue of satisfying one of your desires, it claims that the kind *thing that satisfies one of your desires* is the only basic good.<sup>4</sup>

A *welfare subject* is an entity that is capable of being positive or negative in welfare. It might seem that any theory of welfare must tell us which entities are the welfare subjects. This requirement should be rejected, however, because it stacks the deck against variabilism. If desire satisfactionism is true of all welfare subjects, then an entity is a welfare subject if and only if it is capable of desire. But the view that desire satisfactionism is true of humans—a view that admits that other theories may be true of other entities—has no such implication: it is compatible, for example, with the claim that plants are welfare subjects even though they cannot have desires. If invariabilism is true, then the one true theory of welfare enables us to identify all of the welfare subjects. But if variabilism is true, we can do this only if we have *all* of the (possibly many) true theories. To avoid begging the question against variabilism, we must not require every theory to tell us which entities are welfare subjects.

In light of these facts about the functions that any theory of welfare must perform, we can clarify what’s at issue between invariabilism and variabilism. Invariabilism does not imply that the same *particulars* are basically good for all welfare subjects: obviously, the claim that (say) hedonism is true of all subjects doesn’t entail that the pleasure that I felt just now benefited all of us. Invariabilism implies that the same *kinds* are basically good or bad for all welfare subjects—i.e., that the same list of basic goods and bads applies to them. Since a theory of welfare must also enable us to calculate the basic prudential values of the tokens of the basic goods and bads, invariabilism also says that the same story about how to calculate these values applies to all subjects. By contrast, variabilism implies *either* that no list of basic goods and bads applies to all subjects *or* that the basic prudential values of the tokens of at least one basic good or bad are calculated differently for different subjects.<sup>5, 6</sup>

---

<sup>4</sup> I elaborate on this in Lin (2016b).

<sup>5</sup> I assume that for every welfare subject, some theory of welfare is true of that subject. Thus, I take variabilism to imply that at least one theory of welfare is true of one subject but false of another.

<sup>6</sup> Fletcher (2009) uses the terms ‘variabilism’ and ‘invariabilism’ to draw a distinction that is orthogonal to the one with which I am concerned. On his terminology, invariabilism affirms (whereas variabilism denies) that if a kind *K* is a basic good for a subject *S*, then for *every* member *x* of *K*, *S*’s getting *x* would be basically good for *S*—regardless of the context in which this would occur. Invariabilism in Fletcher’s sense is compatible with both variabilism and invariabilism in my sense. Like nearly all theorists, whether variabilist or invariabilist, I assume that the view that Fletcher calls ‘invariabilism’ is true.

## 2. The Case for Invariabilism

I will now argue that we should accept invariabilism. But before I do, let me mention a way of arguing for invariabilism that I will not be invoking. The contrast between my strategy and the one that I am setting aside will clarify what is distinctive about my approach.

Suppose that all and only sentient beings are capable of welfare, and that hedonism is the correct theory of the welfare of cognitively typical human adults. Given these assumptions, invariabilism is difficult to resist—at least if there are no creatures more advanced than cognitively typical human adults. For if hedonism is true of *us*, then none of the kinds besides pleasure that might be thought to be basic goods for us (e.g., achievement, knowledge, friendship) are in fact basic goods for us. And if none of those kinds is a basic good for us, it is implausible that any of them is a basic good for more primitive welfare subjects, like dogs or cats. Thus, if the aforementioned assumptions are true, then pleasure is surely the only basic good for all subjects, and for similar reasons, pain is surely the only basic bad for all subjects. But if this is so, then hedonism is most likely true of all subjects. For whatever story hedonism tells about the basic prudential values of pleasures and pains felt by us, the same story will presumably apply to any other creatures for whom pleasure and pain are the only basic good and bad.

The argument that I just sketched exemplifies the strategy of arguing for invariabilism via the claim that a particular theory of welfare is true of us. Since I am not aware of any decisive arguments for any theory of our welfare whose truth would strongly support invariabilism, I will not employ this strategy. Besides, I hope to show that invariabilism can help us in the search for the correct theory of our welfare by ruling out certain theories while making others more attractive. I cannot do this if I assume that hedonism, or some other particular theory, is true of us.<sup>7</sup>

I will argue that there are two main reasons to favor invariabilism. The first can be stated briefly, but the second will require a great deal of discussion.

---

<sup>7</sup> Another argument for invariabilism that I will not explore is given by McDaniel (unpublished), who assumes that prudential value is a form of *intrinsic* value. While there is much to admire in McDaniel’s paper, I prefer to argue for invariabilism without relying on this assumption.

## 2.1 THE SIMPLICITY OF INVARIABILISM

The first reason to favor invariabilism is its simplicity. According to variabilism, at least one theory is true of some subjects but false of others, and there could even turn out to be a large plurality of true theories, each of them true of different subjects. By contrast, according to invariabilism, welfare is “one size fits all.” This is a simpler picture than the one available on variabilism, and the greater simplicity of a view is a reason to favor it.

## 2.2 THE INEXPLICABILITY OF VARIABILISM

The second and more important reason to favor invariabilism is that it is difficult, on reflection, to see what could explain why a theory of welfare might be true of some subjects but false of others. Admittedly, it is natural to suppose that the many differences in the natures or capacities of welfare subjects could explain why different theories might be true of them. But as I will now argue, this impression is mistaken.

Let’s begin with the most obvious way a theory that is true of some subjects might be false of others: it might identify as a basic good a kind that *isn’t* a basic good for the latter subjects.<sup>8</sup> There are two *prima facie* plausible strategies for explaining why a given kind might be a basic good for some subjects but not others. I will argue that both strategies are implausible on close inspection.

### *2.2.1 The Inaccessibility Strategy*

I call the first of these the Inaccessibility Strategy. It begins with the observation that, among the kinds that *might* turn out to be basic goods for us, there are kinds that are *inaccessible* to certain welfare subjects, in the sense that those subjects lack the physical or psychological capacities that are needed to possess, or to otherwise be suitably related to, tokens of those kinds. *Theoretical contemplation* might be a basic good for us, but it is inaccessible to dogs, at least as they actually are: given how they are actually constituted, dogs are incapable of engaging in it.<sup>9</sup> Now, if a kind is inaccessible to a given

---

<sup>8</sup> Another possibility, which I will return to later, is that although the theory correctly lists the basic goods and bads for the latter subjects, it incorrectly calculates the basic prudential values of the tokens of some of those basic goods or bads.

<sup>9</sup> Note that the relevant sense of capacity is more robust than mere possibility: even if there are distant possible worlds in which dogs engage in theoretical contemplation, they are actually incapable of engaging in it.

subject, then it is natural to suppose that it is not a basic good for it. Thus, it seems easy to explain why a basic good for some subjects might lack that status for others: it might be accessible only to the former subjects. For example, the fact that Fido cannot engage in theoretical contemplation seems to explain why it is not a basic good for him even if it turns out to be a basic good for us.

The Inaccessibility Strategy fails, however, because it relies on the following false assumption:

**Inaccessibility Excludes Goodness**      If a subject *S* is incapable of possessing (or otherwise being suitably related to) tokens of a kind *K*, then *K* is not a basic good for *S*.

To see why this assumption is false, we need to get clearer on what is implied by the claim that a given kind is a basic good for certain welfare subjects. Consider Anhedonic Annie, a human being who lacks the physiological and psychological capacities that are required for pleasure. In spite of this unfortunate fact about her, pleasure could still be a basic good for her. The claim that pleasure is a basic good for her doesn’t imply that she ever actually feels any pleasure. Nor does it imply that she feels any pleasure at any nearby possible worlds. It merely implies that for any pleasures, if she *were* to experience them, each of them *would* be basically good for her. And this counterfactual could be true even though Annie is incapable of feeling pleasure.

An analogy might be helpful here. Rossian pluralism says that lying is *prima facie* wrong. Suppose that there’s an agent who is incapable of lying. It doesn’t follow from this that lying isn’t *prima facie* wrong for this agent, and hence that Rossian pluralism is false of him. Rossian pluralism merely implies that, for any acts of lying, if this agent *were* to perform those acts, each of them would be *prima facie* wrong. Lying—the kind—can be *prima facie* wrong for an agent, even if the agent is incapable of performing any action of that kind. Likewise, pleasure—the kind—can be a basic good for Annie even if she is incapable of possessing any member of that kind.

More generally, the claim that *K* is a basic good for a certain subject, *S*, merely implies the following about the possession by *S* of tokens of *K*: for any tokens of *K*, if *S* were to have them, each of them would be basically good for *S*.<sup>10</sup> This counterfactual can be true even if *K* is inaccessible to *S*. Even

---

<sup>10</sup> There are two necessity operators implicit here. We are concerned with all *possible* tokens of *K*, and the counterfactual is meant to be true at all possible worlds: necessarily, for any *xx*, if each of *xx* is a token of *K*, then necessarily, if *S* were

though dogs are incapable of theoretical contemplation, it could be true that for any instances of theoretical contemplation and any dog  $S$ , if  $S$  were to engage in them, each of them *would* be basically good for  $S$ . Thus, theoretical contemplation could be a basic good for Fido, even though it is inaccessible to him. Inaccessibility Excludes Goodness is false, so the Inaccessibility Strategy fails.

We should think about basic goods in this way because ethical theory should be fully general. A theory of right action should be able to tell us which actions are right in any situation, not just in nearby possible worlds. Likewise, a theory of a subject  $S$ 's welfare should be able to tell us how well off  $S$  is in any situation—even ones in which  $S$  possesses tokens of a kind that is inaccessible to it. Since it's epistemically possible that Fido would intrinsically benefit from engaging in theoretical contemplation if he were to do so, the fact that he cannot actually do so shouldn't exclude it from being a basic good for him: after all, it might be relevant to his welfare in counterfactual scenarios in which he engages in it. To ensure that our theories of welfare have the requisite level of generality—to ensure that they can tell us how well off the relevant subjects are in *any* situation—we must reject Inaccessibility Excludes Goodness.

Admittedly, the claim that theoretical contemplation is a basic good for Fido sounds odd. But this is because it would be *misleading* to enumerate it when asked for a list of his basic goods. Although a theory of Fido's welfare should be fully general, a request for such a list is naturally interpreted as concerning the basic goods that fix how well off he is at *nearby* possible worlds. A kind cannot help fix a subject's level of welfare at worlds at which it is inaccessible to that subject: if  $S$  cannot possess any tokens of  $K$  at  $W$ , then no tokens of  $K$  are basically good or bad for  $S$  at  $W$ . Thus, to enumerate theoretical contemplation in response to such a request might suggest that, as he actually is, Fido can engage in such contemplation. None of this disproves the claim that for any instances of theoretical contemplation, if Fido *were* to engage in them, each of them would be basically good for him. A kind can be a basic good for a subject even though its inaccessibility to him means that we can ignore it for practical purposes (e.g., when trying to increase his welfare). We cannot give a subject tokens of a kind that is inaccessible to it, so we cannot promote a subject's welfare by doing this. But although the inaccessibility of a kind excludes its practical relevance, it doesn't exclude its being a basic good.

---

to have  $xx$ , each of  $xx$  would be basically good for  $S$ . Moreover, as I explain in Lin (2016b), and as I alluded to earlier, the claim that  $K$  is a basic good for  $S$  also implies the existence of a good-making property. But this is not relevant for present purposes.

It bears emphasizing that the claim that theoretical contemplation is a basic good for Fido does not imply that Fido actually *benefits* from such contemplation. The only way to benefit from a basic good is to possess (or to otherwise be suitably related to) some token of it. Since Fido never engages in theoretical contemplation, he never benefits from it. The claim that such contemplation is a basic good for him merely implies that he *would* benefit from any instances of it if he were to engage in them. These merely counterfactual benefits have no effect on Fido’s actual level of welfare. The distinction between basically good *kinds* (i.e., basic goods) and basically good *particulars* must be kept in mind. If *S* possesses a basically good particular *x*, then *S* benefits from *x*. But *S* might not actually benefit from any member of *K* even if *K* is a basic good for *S*, because *S* might not actually possess any member of *K*. Let me also remind you that I am not claiming that theoretical contemplation *is* a basic good for Fido, or indeed, for any other subject. My point is merely that, from the fact that it is inaccessible to Fido, it doesn’t follow that it isn’t a basic good for him.

You might worry that my view of basic goods falls prey to the conditional fallacy. On the view that I have proposed, pleasure is a basic good for me only if, for any pleasure *x*, if I were to feel *x*, *x* would be basically good for me. But surely, there is a possible scenario in which feeling a particular pleasure causes me to die painfully. Does this show that the relevant counterfactual is false, and that I must therefore deny that pleasure is a basic good for me? No. How well off I would be if I were to have *x* depends on how basically good or bad for me *everything* would be in that scenario. It could be that *x* would be basically good for me, even though its basic goodness would be outweighed by the basic badness of other things, including its effects. Once we distinguish (i) how basically good or bad for me *x* would be if I were to have it, from (ii) what my overall level of welfare would be if I were to have *x*, the worry disappears.<sup>11</sup>

Another objection might suggest itself. You might think that theoretical contemplation is not merely inaccessible to Fido, but *necessarily* so: not only is he incapable of engaging in it as he actually is, but there are no possible worlds at which he is capable of doing so. If this is true, then according to the standard, possible-worlds account of counterfactuals, it is *vacuously* true that for any instances of theoretical contemplation, if Fido were to engage in them, each of them would be basically good for him. Thus, my view seems to imply that theoretical contemplation is a basic good for Fido precisely because it’s necessarily inaccessible to him. Indeed, it seems to imply that if any kind is necessarily

---

<sup>11</sup> See Lin (unpublished) and Heathwood (2005).

inaccessible to a subject  $S$ , then it is a basic good for  $S$ . I have argued that inaccessibility doesn't *exclude* goodness, but I seem saddled with the view that *necessary* inaccessibility *entails* goodness.

I have two replies. First, I doubt that theoretical contemplation really is necessarily inaccessible to Fido. Some would claim that the essence of dogs excludes such contemplation: there is no possible world,  $W$ , and subject,  $S$ , such that  $S$  is a dog at  $W$  and  $S$  engages in such contemplation at  $W$ . But as many other philosophers (e.g., Kagan, Tooley, and McMahan) would agree, this claim is dubious.<sup>12</sup> Moreover, even if it is true, it does not follow that there is no possible world at which *Fido* engages in such contemplation. For perhaps Fido is not *necessarily* a dog: there may be possible worlds at which he engages in theoretical contemplation *and is not a dog*. Never mind whether Fido would still be a dog if he were to engage in theoretical contemplation. Is it really so obvious that no amount of technology or wizardry could endow him with the capacity to contemplate?<sup>13</sup>

But even if we grant that theoretical contemplation is necessarily inaccessible to Fido, the objection can be answered. The objection assumes the standard, possible-worlds account of counterfactuals, on which a counterfactual is true just if all of the closest *possible* worlds at which its antecedent is true are ones where its consequent is true. But as many philosophers now recognize, we should reject or amend this view precisely because it implies that every counterfactual with an impossible antecedent is true. After all, some counterfactuals with impossible antecedents are intuitively false, or at least not true (e.g., “If 323 were prime, it would be divisible by 2, 4, and 16”).<sup>14</sup> The simplest emendation of the account that accommodates this fact appeals to impossible worlds: a counterfactual is true just if all of the closest worlds—possible or impossible—at which its antecedent is true are ones where its consequent is true.<sup>15</sup> But we needn't endorse this account to recognize that the correct theory of counterfactuals, whatever it turns out to be, will allow some counterfactuals to be false (or at least untrue) even though they have impossible antecedents.

---

<sup>12</sup> Kagan (2016, p. 13) imagines enhancing a dog's cognitive capacities until it is a person, and he refers to the resulting creature as an “intelligent dog.” Tooley (1972, p. 60-62) claims that a cat could acquire “all the psychological capabilities characteristic of adult humans” while remaining a cat. McMahan (2002, pp. 147-49) says that a chimpanzee could, while remaining a chimpanzee, be endowed with the psychological capacities of a 10-year-old human child.

<sup>13</sup> Kagan, Tooley, and McMahan all assume that the individuals they describe would survive the relevant transformations.

<sup>14</sup> Nolan (2013), p. 363.

<sup>15</sup> Nolan (2013), p. 363.

Thus, the present objection is merely an artifact of a well-known limitation of the possible-worlds account of counterfactuals. Because not all counterfactuals with impossible antecedents are true, the aforementioned counterfactual involving Fido and theoretical contemplation is not vacuously true—even if such contemplation is necessarily inaccessible to Fido. Thus, my account does not imply that theoretical contemplation is a basic good for Fido. Nor does it imply that any kind that is necessarily inaccessible to a subject is therefore a basic good for that subject. Necessary inaccessibility does not entail goodness.

Notice, though, that the relevant counterfactual about Fido could still be true, even though it isn’t vacuously true: it could be true that for any instances of theoretical contemplation, if Fido were to engage in them, each of them would be basically good for him.<sup>16</sup> Thus, theoretical contemplation could be a basic good for Fido, even if it is necessarily inaccessible to him. Necessary inaccessibility does not *exclude* goodness: *K* can be a basic good for *S* even though there are no possible worlds at which *S* is capable of possessing any tokens of *K*. For even if there are no such worlds, it could be that for any tokens of *K*, if *S* were to possess them, each of them would be basically good for *S*.

You might worry that if necessary inaccessibility doesn’t exclude goodness, then we can’t identify the basic goods (if any) that are necessarily inaccessible to a subject: if *K* is necessarily inaccessible to *S*, how can we tell whether it’s true that for any tokens of *K*, if *S* were to possess them, each of them would be basically good for *S*? But this worry is unfounded. Our intuitions about the truth values of the relevant counterfactuals remain unchanged whether we consider their antecedents possible or impossible. It is an equally open question whether the aforementioned counterfactual about Fido is true, regardless of whether its antecedent is possible. By contrast, “if Fido were despondent about his middling grade on a mathematics exam, this despondency would be basically good for him” is intuitively false, whether or not such despondency is necessarily inaccessible to Fido. When we attempt to evaluate such counterfactuals, it doesn’t matter whether their antecedents are possible or impossible. Thus, no special problems are introduced by the view that necessary inaccessibility doesn’t exclude goodness: we can use whatever method we would otherwise use (e.g., reflective equilibrium) to determine which kinds are basic goods.

---

<sup>16</sup> Because this counterfactual concerns a scenario in which the laws of nature may be different, an editor at *Ethics* doubts that our intuitions about it are reliable. But while a difference in the laws might affect whether contemplation has good effects, it seems to me that it couldn’t affect whether instances of it are *basically* good for the subjects engaging in them.

To summarize, here is my argument against the Inaccessibility Strategy. The claim that a kind *K* is a basic good for a subject *S* merely implies the following concerning the possession by *S* of tokens of *K*: for any tokens of *K*, if *S* were to have them, each of them would be basically good for *S*. Because this counterfactual could be true even though *K* is inaccessible (or even necessarily inaccessible) to *S*, Inaccessibility Excludes Goodness is false. Thus, the Inaccessibility Strategy fails: you cannot explain why a basic good for some subjects might fail to be a basic good for others simply by noting that the kind in question might be inaccessible to the latter subjects.

### 2.2.2 *The Perfectionist Strategy*

There is a second *prima facie* plausible strategy for explaining why a basic good for some subjects might fail to be a basic good for others: the Perfectionist Strategy. According to this strategy, the *nature* of a subject constrains which kinds can be basic goods for it—even if, as I have argued, a kind whose tokens are inaccessible to a subject can nonetheless be a basic good for that subject. As Richard Kraut puts it:

When we assess whether *G* is good for *S*, we must assess both *G* and *S* to see whether they are properly matched to each other.... [T]he “for” in “*G* is good for *S*” is best taken to indicate that *G* has a certain kind of suitability to *S*....<sup>17</sup>

The Perfectionist Strategy grants that inaccessibility doesn’t exclude goodness, and it accepts the counterfactual-based account of basic goods that I proposed. However, it says that whether the relevant counterfactual is true of a given subject depends on whether the putative basic good is well-suited to its nature. Suppose that theoretical contemplation is a basic good for us. Perhaps the fact that we are *rational* explains why it’s true that for any instances of theoretical contemplation, if any of us were to engage in them, each of them would be basically good for him. And perhaps the fact that Fido isn’t rational explains why the corresponding counterfactual is false of him.

But this is unpersuasive. If we are rational creatures, this is because we possess certain rational capacities, including the capacity for theoretical contemplation. Fido actually lacks those capacities, so he isn’t actually rational. But if he *were* to engage in theoretical contemplation, he *would* possess

---

<sup>17</sup> Kraut (2007), pp. 85, 86-87. Also see Kraut (1994), pp. 47-48.

those capacities, so he *would* be rational. Any scenario in which Fido engages in such contemplation is one in which a rational creature engages in it, and thus one in which a creature well-suited to it engages in it. Why deny, then, that theoretical contemplation is a basic good for Fido if it is a basic good for us? Why insist that Fido’s *actual* nature determines what he would benefit from in far-out *counterfactual* scenarios in which he has a very different nature?

Remember that the claim that theoretical contemplation is a basic good for Fido doesn’t imply that he ever actually benefits from such contemplation: it only implies that he benefits from it in the counterfactual scenarios in which he engages in it. Thus, on the view that theoretical contemplation is a basic good for Fido, there is never a mismatch between Fido’s nature and what he’s benefiting from. At the actual world, Fido isn’t rational, and he isn’t benefiting from theoretical contemplation. In the counterfactual scenarios in which he engages in it and benefits from it, he is rational. The fact that he isn’t *actually* rational shouldn’t preclude him from benefiting from it if he *were* rational.

It might be replied that it is worse for a human adult to lack rational capacities than it is for Fido to do so, and that we have more reason to bestow these capacities on a human adult who lacks them than on Fido. But these claims do not support the Perfectionist Strategy. We are inclined to accept them, if we are, because we find it plausible that typical human lives are as good as they are at least partly on account of containing exercises of rational capacities. Thus, we deem it especially bad for a human adult to lack such capacities, in the sense that because he lacks them, his welfare is much lower than is typical for human adults.<sup>18</sup> By contrast, a dog who lacks such capacities is not therefore significantly lower in welfare than dogs typically are. If we have especially strong reasons to bestow rational capacities on a human adult who lacks them, that is because we have very strong reasons to ensure that no subject is much lower in welfare than members of its kind typically are.<sup>19</sup> None of this undermines the claim that for any tokens of any kind that is a basic good for us, if Fido were to possess those tokens, each of them would be basically good for him. Thus, none of it explains why a basic good for us might fail to be a basic good for Fido.

---

<sup>18</sup> If, like Campbell and Stramondo (forthcoming), you deny that human adults who lack rational capacities are much worse off than those who possess them, then you won’t be drawn to the present objection anyway.

<sup>19</sup> McMahan (2002, pp. 144-65) contrasts a subject’s level of welfare with how *fortunate* it is—where the latter is largely determined by where its level of welfare is within the range of welfare available to subjects of its kind, as well as by how that level compares to those of other members of its kind. A dog who lacks rational capacities is not *unfortunate*, but a human adult who lacks them may well be. This may be why we have more reason to bestow such capacities on a human adult who lacks them than we do to bestow them on a dog.

One might appeal to intuitions about other cases to support the view that differences in natures can explain differences in basic goods. One might claim, for example, that eating grass is a basic good for cows even though it isn't a basic good for us. But intuitions of this sort conflate derivative and basic goodness. It is derivatively good for cows to eat grass, since they are well-suited to digest it, and eating something that one is well-suited to digest is conducive to good health. But there is no reason to think that eating grass is *basically* good for cows. The fact that it is good for cows to eat grass suggests, at best, that health is a basic good for cows. But if this is a basic good for cows, then there is no reason to suppose that it isn't also a basic good for us: we are just as well-suited to being healthy as cows are. Something similar could be said about other intuitions that appear to support the Perfectionist Strategy.

Finally, notice that one cannot explain why a basic good for some subjects might fail to be a basic good for others by endorsing perfectionism for all subjects. Perfectionism says that every particular that is basically good for a subject has that status *because* it is an exercise of a capacity that is essential to that subject.<sup>20</sup> Even if it identifies a plurality of capacities the exercise of which is basically good for a given subject, it claims that there is only one property that can explain why certain particulars are basically good for that subject: *being an exercise of a capacity that is essential to that subject*. Since, as I argued earlier, each basic good corresponds to a good-making property, perfectionism claims that the exercise of one's essential capacities is the sole basic good.<sup>21</sup> Thus, if perfectionism is true of all welfare subjects, then even if different capacities are essential to different subjects, the exercise of one's essential capacities is the sole basic good for all subjects. Even if rationality is essential to humans but not dogs, what explains why any exercise of rationality is basically good for a human is the fact that it is an exercise of one of that subject's essential capacities. This is also what explains why any exercise of olfaction (assuming that it is an essential canine capacity) is basically good for a dog. Thus, if perfectionism is true across the board, then the same single-entry list of basic goods applies to all welfare subjects. This should not be surprising. Perfectionism is a theory of welfare, so the claim that it is true of all welfare subjects cannot support the hypothesis that different theories are true of different subjects.

---

<sup>20</sup> See, e.g., Bradford (2016). Other versions of perfectionism focus on different capacities, but my remarks could easily be adapted to those versions of the view.

<sup>21</sup> For a fuller defense of this view about good-making properties and basic goods, see Lin (2016b).

### 2.2.3 Basic Prudential Values

I have argued that the Inaccessibility and Perfectionist strategies fail. Besides these strategies, there appear to be no *prima facie* plausible ways to explain why a basic good for some subjects might fail to be a basic good for others. Thus, it is difficult to see how a list of basic goods for one subject might fail to apply to other subjects. Because similar considerations apply to basic *bads*, it is difficult to see how the same list of basic goods and bads could fail to be true of all welfare subjects.

The claim that the same list applies to all subjects is not yet invariabilism, though. As I explained earlier, a theory of welfare must tell us how to calculate the basic prudential value of any token of a basic good or bad (e.g., *how* basically good or bad for its subject any particular pleasure is, if pleasure is a basic good). Even if the same list of basic goods and bads is true of all welfare subjects, different theories could be true of different subjects because the basic prudential values of the tokens of some basic good or bad are fixed differently for different subjects.

Although there is nothing incoherent about this hypothesis, we should reject it on the grounds that its truth cannot plausibly be explained. Suppose that pleasure is a basic good for all welfare subjects. If the value of a pleasure for a dog is just the product of its intensity and duration, for example, then why would the value of a pleasure for a human be determined by some other function?

It might be suggested that since humans are capable of feeling *higher* pleasures, the basic prudential value of a pleasure for a human is the product of its intensity, duration, and *quality*. But if this is true, then surely quality is also relevant in the same way to the values of canine pleasures. Perhaps dogs are actually capable of feeling only the lowest grade of pleasure. But if any dog were to feel a higher pleasure, the value of this pleasure would be amplified in proportion to its quality. Thus, if we accept the present suggestion about human pleasures, we should also believe that the value of a pleasure for a dog is the product of its intensity, duration, and quality—where the last of these factors always has the same low value (e.g., 1) when actual dogs are in question.

Alternatively, it might be suggested that facts about the *shape* of a life affect the values of pleasures in humans but not in dogs, since humans (unlike dogs) are capable of caring about the shapes of their lives, or capable of engaging in the sorts of goal-directed activities that can give a life an evaluatively

significant shape. (For example, it might be that when a human’s pleasure is narratively connected in the right ways with her goals or past efforts, its value is amplified.)<sup>22</sup> But as before, we should ask what would be true of dogs if they had the relevant capacities. If the values of human pleasures are affected in this way because humans have those capacities, then surely, the values of canine pleasures would also be affected in the same way if dogs had those capacities. Thus, if we think that the shape of a life affects the values of human pleasures, we should think that it affects the values of canine pleasures in the same way, but that this factor can be ignored when actual dogs are in question.<sup>23</sup>

As the foregoing examples show, the sorts of argument that I gave against the Inaccessibility and Perfectionist Strategies can also be leveled against attempts to explain why the values of the tokens of a given basic good or bad might be calculated differently for different subjects. Because I see no plausible explanation of why this might be so, I claim that if the same list of basic goods and bads applies to all welfare subjects, then the same theory of welfare is true of all welfare subjects. But as I argued earlier, it is difficult to see how the same list of basic goods and bads might fail to apply to all welfare subjects. Thus, it is difficult to see how the same theory might fail to be true of all welfare subjects. This concludes my case for the second main reason to favor invariabilism: that it is difficult to see what could explain the truth of variability.

### 2.3 SUMMARY OF THE CASE FOR INVARIABILISM

To summarize, here is why we should accept invariabilism. There are two main reasons to favor invariabilism. First, it is simpler than variability. Second, it is difficult to see what could explain the truth of variability. Of course, these reasons must be weighed against any reasons there might be to favor variability. But there appear to be no significant reasons to favor variability. If differences in the natures and capacities of welfare subjects cannot explain why a theory might be true of some subjects but false of others, then I see no motivation for variability. Because there are two strong reasons to favor invariabilism and there appear to be no significant reasons to favor variability, I conclude that we should accept invariabilism.

---

<sup>22</sup> I thank an anonymous reviewer for this suggestion. For more on these issues, see Velleman (1991), Kauppinen (2012), and Dorsey (2015).

<sup>23</sup> Relatedly, perhaps some facts about the shape of a life (e.g., that it improves over time) are *themselves* basically good for humans. But in light of my earlier arguments, it is hard to see why such facts would not also be basically good for dogs.

### 3. Implications

Invariabilism doesn’t commit us to any particular theory of welfare. It merely commits us to thinking that a given theory is true of *some* subject if and only if it is true of *all* subjects. Nevertheless, my argument has implications for the question of which theory is true of us: it eliminates many of the going theories while making some of the remaining ones more attractive.

#### 3.1 OBJECTIVE LIST THEORIES

An *objective list theory* of welfare is one on which there are a plurality of basic goods, at least some of which are independent of the subject’s favorable attitudes.<sup>24</sup> Elsewhere, I have expressed sympathy for the view that the correct theory of our welfare is an objective list theory.<sup>25</sup> But such theories face a problem—or, at least, many of them do. Many objective list theories posit some basic goods that are either inaccessible to certain welfare subjects or not properly matched to their actual natures (e.g., theoretical contemplation). One apparent disadvantage of such a theory is that even if it is true of us, it cannot be true of those subjects. Accepting such a theory of our welfare seems to commit us to a second kind of undesirable complexity, besides pluralism about the basic goods for us: pluralism about the number of true theories of welfare.

But in light of what I have argued, it should be obvious that this appearance is misleading. A theory of this kind could be true of all welfare subjects, since a kind can be a basic good for a subject even though it is inaccessible to it, not properly matched to its actual nature, or both. If the correct theory of our welfare is an objective list theory on which theoretical contemplation is a basic good, then this theory could also be the correct theory of Fido’s welfare. Thus, my argument for invariabilism neutralizes an important objection to many objective list theories.

#### 3.2 SOPHISTICATED THEORIES OF WELFARE

Many theories of welfare differ with respect to the cognitive or psychological capacities that they require on the part of the subjects to which they are meant to apply. On some theories, a particular

---

<sup>24</sup> For a more precise characterization, see Lin (2016a).

<sup>25</sup> Lin (2014) and (2017).

is basically good for any subject of the relevant kind only if it is related to that subject in a way that requires a fair amount of cognitive or psychological sophistication on the part of that subject. We can call such theories *sophisticated* theories—keeping in mind that a theory might be intricate or complex in its structure (e.g., in the number of basic goods it postulates, in how it fixes the basic prudential values of particulars, etc.) without being sophisticated in this sense.

Sophisticated theories of the welfare of typical human adults have been on the rise in recent years. One example of such a theory is the *Person-Centered Theory*, according to which the sole basic good is the realization of your values. Following Scheffler, this view maintains that valuing X involves four conditions: (i) a belief that X is good, (ii) a disposition “to experience a range of context-dependent emotions” concerning X, (iii) a disposition to experience these emotions as warranted, and (iv) a disposition “to treat certain kinds of X-related considerations as reasons for action.” On this view, a particular can be basically good for you only if you possess the capacity to value something. This capacity requires a fair degree of sophistication, since it involves the capacities to have evaluative beliefs and to treat certain considerations as reasons for action.<sup>26</sup>

Although I won’t go into details here, a number of other sophisticated theories have been proposed. These include Dorsey’s *Judgment Subjectivism*, Rosati’s *Successful Loving* view, Tiberius’s *Value Fulfillment Theory*, Bruckner’s *Minimal Accountability Desire Satisfactionism*, and Sumner’s *Authentic Happiness* view.<sup>27</sup>

Sophisticated theories can be contrasted with *simple* theories, which do not require sophisticated cognitive or psychological capacities on the part of all the subjects to which they are meant to apply. The most prominent traditional theories are all simple. Hedonism says that some particulars can be basically good for you as long as you have the capacity to feel pleasure. Desire satisfactionism says that some particulars can be basically good for you as long as you have desires. Even an objective list theory on which some basic goods require sophistication can be simple, since it can include pleasure or desire satisfaction on its list of basic goods: since a member of one of these kinds can be basically good for you even if you lack sophisticated capacities, such a view would not say that *every* particular that is basically good for you requires you to have such capacities. Indeed, any sophisticated theory

---

<sup>26</sup> Yelle (2014), quoting from Scheffler (2011).

<sup>27</sup> Dorsey (2012); Rosati (2006); Tiberius (2014); Bruckner (2016); and Sumner (1996), ch. 6. The view defended in Raibley (2010) and (2013) comes close, but it isn’t sophisticated in my sense.

can be turned into a pluralistic simple theory by adding pleasure, desire satisfaction, or some other psychologically simple kind to its list of basic goods—a point to which I will return later.<sup>28</sup>

Sophisticated theories are false of the many welfare subjects who lack sophisticated capacities (e.g., newborns and dogs). The claim that a sophisticated theory is true of such a subject implies that no particulars are basically good for it, and hence that it is not positive in welfare. But clearly, many subjects who lack sophisticated capacities are nonetheless positive in welfare: many newborns and dogs are plainly faring well. Thus, sophisticated theories are false of these subjects.<sup>29</sup>

Notice that this argument does not appeal to the discredited principle that Inaccessibility Excludes Goodness. My claim is not that a sophisticated kind cannot be *a* basic good for a simple subject: as I said earlier, it can. My claim is that such a kind cannot be the *only* basic good for a simple subject *who is actually positive in welfare*. Theoretical contemplation can be among the basic goods for Fido, but it cannot be the *only* basic good for him if he is actually positive in welfare. A subject *S* is positive in welfare only if some particulars are basically good for it. A particular is basically good for *S* only if it belongs to a kind whose tokens *S* has the capacity to possess—i.e., one that is accessible to *S*. Thus, since many newborns and dogs are positive in welfare, at least one basic good is accessible to them. Sophisticated theories imply that *no* basic good is accessible to them, and hence that none of them is positive in welfare. This is why sophisticated theories are false of simple subjects.

One might grant that sophisticated theories aren’t true of all welfare subjects but claim that some such theory is nonetheless true of us. After all, we possess the capacities that these views invoke, so they don’t imply that *we* aren’t positive in welfare. But if invariabilism is true, any theory that is false of some subjects is false of all subjects. Since all sophisticated theories are false of some subjects, invariabilism entails that all such theories are false of us. Indeed, it entails that they are false, period.

---

<sup>28</sup> What about a view on which *x* is basically good for *S* just if an idealized counterpart of *S* (e.g., *S* as he would be if he were fully informed and rational) has a sophisticated attitude toward *x* at a different possible world, *W*? Such a view is simple rather than sophisticated, since it could be true that *S*’s idealized counterpart has such an attitude toward *x* at *W* even though *S* lacks sophisticated capacities. I thank an anonymous reviewer for raising this issue.

<sup>29</sup> An anonymous reviewer imagines a view according to which the sole basic good is the realization of one’s values, but what counts as valuing for a given subject depends on how sophisticated that subject is: valuing in typical human adults requires the exercise of sophisticated capacities, but for simple creatures like newborns, desiring is sufficient for valuing. This view evades the present argument, since subjects who lack sophisticated capacities could nonetheless be positive in welfare if this view were true of them. But precisely because it does not require sophisticated capacities on the part of all of the subjects to which it is meant to apply, this view is simple rather than sophisticated.

Thus, if invariabilism is true, we can dramatically narrow down the range of theories of welfare that might turn out to be true: the only live options are theories that are *simple* in the sense I explained.<sup>30</sup>

#### 4. Objections and Replies

Let me end by considering some objections.

##### 4.1 THE SOPHISTICATED THEORIST’S REPLY

Earlier, I claimed that there seem to be no significant reasons to favor variability. But invariabilism allows us to rule out all sophisticated theories of welfare. Those who defend such theories will deem this very fact a reason to accept variability. Since there is surely *some* reason to think that the correct theory of our welfare is a sophisticated one, this gives us some reason to accept variability. But as I will now argue, this reason is much weaker than the ones that I gave in support of invariabilism.

As I said earlier, any sophisticated theory can be turned into a simple theory by adding pleasure, desire satisfaction, or some other psychologically simple kind to its list of basic goods. For example, the Person-Centered Theory (PCT) would be a simple theory if it were altered to say that *two* kinds are basic goods: the realization of your values, and pleasure. Nearly all of the attractive features of sophisticated theories can be captured by simple pluralistic views of this kind, on which one or more of the basic goods involve the sophisticated capacities that those theories invoke. Thus, when it comes to the welfare of typical adults, we have very little reason to accept a sophisticated theory instead of accepting a closely related simple theory.

Consider the PCT and the aforementioned modification of it. One main virtue of the PCT is that, unlike hedonism and desire satisfactionism, it need not say that you are very high in welfare if you spend your life taking a pleasant drug that you intensely and addictively desire. Since you might not *value* the drug even though you desire and enjoy it, the drug-addicted life might be one in which you hardly realize any of your values, and thus one in which you are low in welfare.<sup>31</sup> But the simple theory that we arrived at by amending PCT can say something similar: if you don’t value the drug,

---

<sup>30</sup> This is also noted in McDaniel (unpublished).

<sup>31</sup> The view defended by Raibley (2010) also has this implication.

then your addiction prevents you from realizing most of your values, so you are worse off than you would be if you weren’t addicted. This is especially true if the view maintains, as it might, that the realization of your values is given much more weight in determining your welfare than pleasure: the pleasure that you get from the drug might give you much less basic goodness than you would be getting from the valued things that your addiction precludes. This view, which is simple in my sense even though it posits two basic goods, has the obvious virtues of PCT: it says that the realization of your values is the strongest welfare-determining factor, and that a life of mindless pleasure is likely to be low in welfare precisely because it prevents you from realizing your values.<sup>32</sup> I see little reason to prefer PCT over this view, and I claim that something similar would be true of any sophisticated theory. Thus, we have little reason to think that a sophisticated theory is true of us, and the fact that invariabilism rules out such theories gives us little reason to reject it.

Indeed, I have argued elsewhere that we should think that sophisticated theories are false of us— even if variabilism is true.<sup>33</sup> If this is so, then we have even less reason to doubt invariabilism on the grounds that it lets us rule out sophisticated theories. My argument begins with the aforementioned observation that sophisticated theories are false of newborns infants. If such a theory is nonetheless true of typical human adults, then it surely *becomes* true of humans when they acquire the capacities that the theory deems necessary for well-being. But the idea that such a theory could go from being false to being true of the same subject has implausible implications. I will illustrate this using PCT, but my remarks generalize to all sophisticated theories.

Recall that according to PCT, a particular is basically good for you just if it is the realization of one of your values. Since many newborns are positive in welfare even though they do not value anything, PCT is false of newborns. Now, consider a human being who acquires the capacity to value things at the normal stage in her development, but who fails to exercise this capacity for some time after she acquires it. Imagine that, prior to acquiring the capacity, she was high in welfare. For the sake of specificity, imagine that hedonism is true of humans who are too young to value things, and that our subject had a very favorable balance of pleasure over pain prior to acquiring the relevant capacity. Furthermore, suppose that the conditions in virtue of which she was high in welfare persist after she acquires the capacity: she retains just as favorable a balance of pleasure over pain after acquiring it. If

---

<sup>32</sup> If the proponent of PCT insists that no drug-addicted life, however euphoric or long-lasting, could be better than any life in which you realize your values, the simple theory could say that values-realization is lexically superior to pleasure.

<sup>33</sup> Lin (2017). I restrict my attention there to sophisticated *subjectivist* theories, but the arguments generalize.

acquiring that capacity makes PCT true of her, then from the time at which she acquires it until the later time at which she begins to value things, no particulars are basically good for her. This implies that between those two times, her welfare is, at best, zero. This is implausible. After all, none of the facts in virtue of which she was previously high in welfare have changed. A subject surely could not go from being high in welfare to being neutral or negative in welfare simply in virtue of acquiring a capacity. Thus, we should not think that PCT becomes true of humans when they acquire the relevant capacity. Nor should we think that PCT is true of us even though it is false of newborns. Since a similar argument can be made against any sophisticated theory, we have strong reasons— independent of any commitment to invariabilism—to think that no such theory is true of us.<sup>34, 35</sup>

I have argued that we have little reason to believe that any sophisticated theory is true of us, and good reasons, independent of invariabilism, to disbelieve this. Such theories therefore give us very little reason to doubt invariabilism. In light of the virtues of invariabilism, we should still accept it.

#### 4.2 THE WELFARE OF PLANTS

Sophistication and simplicity are matters of degree: the capacity to have desires is simple for human beings, but far too sophisticated to be possessed by plants. Couldn't one argue that since many plants are positive in welfare, even simple theories like hedonism and desire satisfactionism are too sophisticated to be true of all subjects? And if so, wouldn't that be a *reductio* of invariabilism?

---

<sup>34</sup> Yelle (2016) replies to this argument by distinguishing a subject's level of welfare *qua* person from her level of welfare *qua* human being and by claiming that PCT is merely a theory of the former. He claims that in the case that I described, the subject continues to fare well as a human being after she acquires the relevant sophisticated capacity: PCT merely says (i) that we can assess her level of welfare *qua* person from the time at which she acquires the capacity, and (ii) that her level of welfare *qua* person after that time is zero. I am not sure that it makes sense to distinguish someone's level of welfare *qua* human being from her level of welfare *qua* person. But even if it does, we need to be told how to determine how well off a subject is overall, or all things considered. If a person's level of welfare *qua* person is given little weight in the calculation of her overall level of welfare, then PCT would be relatively unimportant even if it were true. If not, then PCT has the implausible implication that in the case that I described, the subject's overall level of welfare undergoes a large decline simply because she has acquired the relevant sophisticated capacity. Yelle might claim that there is no such thing as overall welfare, but this is implausible. Surely, it is possible for one person to be better off overall than another.

<sup>35</sup> Dorsey (forthcoming) claims that something is good for a given subject only if the subject stands in a *kinship* relation to it. He argues that although a subject who lacks the capacity to value can stand in such a relation to something even though she doesn't value it, a subject who has this capacity cannot stand in such a relation to anything unless she values it. In his view, this explains why the subject that I described undergoes such a large decline in welfare when she acquires the capacity to value things. I am not moved by this response, however. If it is true that a subject must bear a kinship relation to something in order to benefit from it, then notwithstanding Dorsey's arguments to the contrary, it seems to me that a subject who has the capacity to value things can bear such a relation to something even if she does not value it.

I admit that if any plants are positive in welfare, then invariabilism implies that hedonism and desire satisfactionism are false. In that case, we would have to conclude that something less psychologically demanding than even pleasure or desire satisfaction—something, such as mere life or health, that requires literally no psychological capacities—is a basic good. But I doubt that any plants are positive in welfare, even though they possess tokens of kinds that are *in some sense* good for them. Not every sense of ‘good for’ concerns well-being: when we say that sunlight and water are good for plants, we might merely be making the descriptive claim that they promote the life, growth, or reproductive success of plants.<sup>36</sup> It could be that plants are incapable of possessing any tokens of any basic good or bad, and thus incapable of being positive or negative in welfare.

Even if many plants were positive in welfare, however, this wouldn’t be a *reductio* of invariabilism. Admittedly, invariabilism would rule out hedonism, desire satisfactionism, and any other theory on which a subject must be capable of either pleasure or desire to be positive in welfare. But it would leave it open that the correct theory of welfare is an objective list theory on which some of the goods (e.g., life or health) are simple enough to be accessible to the most primitive welfare subjects while others (e.g., pleasure, desire satisfaction) are accessible only to more sophisticated subjects. Such a view wouldn’t falsely imply that no simple subjects are positive in welfare. Moreover, as I have explained, such a theory could be true of simple subjects even though some of the basic goods it enumerates are inaccessible to them or not well-suited to their actual natures. Thus, a theory of this kind could be correct even if many plants are positive in welfare.<sup>37</sup>

#### 4.3 ARE INANIMATE OBJECTS WELFARE SUBJECTS?

On my view, the claim that *K* is a basic good for *S* merely implies the following concerning the possession by *S* of tokens of *K*: for any tokens of *K*, if *S* were to possess them, each of them would be basically good for *S*. This might be thought to commit me to an implausibly large domain of welfare subjects. Suppose, as is plausible, that pleasure is a basic good on the correct invariabilist theory. In that case, many would find the following counterfactual plausible: for any pleasures and any inanimate object *S*, if *S* were to feel those pleasures, each of them would be basically good for it.

---

<sup>36</sup> See Rosati (2009).

<sup>37</sup> If you balk at the idea that life, health, or any other utterly non-psychological kind is a basic good for us, then you should deny that any plants are positive in welfare. After all, as I have argued, there seems to be no way to explain why any such kind might be a basic good for some subjects but not for others.

If that counterfactual is true, then aren't I committed to the view that pleasure is a basic good for every inanimate object? And wouldn't this have the implausible implication that inanimate objects are welfare subjects?

If the aforementioned counterfactual is true, then I do maintain that pleasure is a basic good for every inanimate object. This view sounds wild, but it is innocuous. Remember that even necessary inaccessibility doesn't exclude goodness:  $K$  can be a basic good for  $S$  even if there are no possible worlds at which  $S$  has any tokens of  $K$ . Thus, the view that pleasure is a basic good for inanimate objects doesn't imply that they are capable of feeling pleasure, or even that there are any possible worlds at which they feel it. *A fortiori*, it doesn't imply that we can make them better off, or that we have reason to benefit them. The view is rendered unobjectionable by the very thing that would commit me to it: my account of basic goods in terms of counterfactuals, together with the fact that counterfactuals with impossible antecedents can be true.

However, from the claim that pleasure is a basic good for inanimate objects, it does not follow that such objects are welfare subjects. An entity is a welfare subject at a world  $W$  just if it is capable at  $W$  of being positive or negative in welfare. Thus, which entities are welfare subjects turns on which theory of welfare is correct, as well as on which entities are capable of possessing tokens of the basic goods and bads identified by the correct theory. If hedonism is true, for example, then an entity is a welfare subject at just those worlds where it is capable of feeling pleasure or pain. It should therefore be evident that, from the view that pleasure is a basic good for inanimate objects, it doesn't follow that any such object is, or even *could* be, a welfare subject. After all, pleasure can be a basic good for such an object even if it is necessarily incapable of feeling pleasure. Thus, even if I maintain that pleasure is a basic good for inanimate objects, this does not commit me to enlarging the domain of welfare subjects. (Similar considerations explain why I need not say that plants are welfare subjects, even if it is true that if any plant were to feel any pleasures, they would be basically good for it.)

Besides, I am not in fact committed to the view that pleasure is a basic good for inanimate objects. For even if I grant that pleasure is a basic good on the correct invariabilist theory, nothing commits me to accepting the aforementioned counterfactual: for any pleasures and any inanimate object  $S$ , if  $S$  were to feel those pleasures, each of them would be basically good for it. More generally, I am not committed to the view that any kind is a basic good for an inanimate object.

#### 4.4 SHALLOW VARIABILISM

As I mentioned above, I reject the following principle:

**Necessary Inaccessibility Excludes Goodness** If a subject  $S$  is such that there are no possible worlds at which it is capable of possessing (or otherwise being suitably related to) tokens of kind a  $K$ , then  $K$  is not a basic good for  $S$ .

After all, on my view, the claim that  $K$  is a basic good for  $S$  merely implies the following concerning the possession of tokens of  $K$  by  $S$ : for any tokens of  $K$ , if  $S$  were to possess them, each of them would be basically good for  $S$ . And this counterfactual can be true even if there are no possible worlds at which  $S$  is capable of possessing any tokens of  $K$ .

Some might hold, however, that  $K$ 's being a basic good for  $S$  doesn't just imply that counterfactual: it also implies that there is a possible world at which  $S$  possesses (and is thus capable of possessing) a token of  $K$ . If this is so, then the aforementioned principle is true, and we can explain why a kind might be a basic good for some subjects but not others: it might be necessarily inaccessible to the latter subjects.

We have good reasons to reject this proposal, however. Consider the kinds that could plausibly turn out to be basic goods for at least some subjects: pleasure, desire satisfaction, friendship, theoretical contemplation, etc. It is hard to know which, if any, of these kinds are necessarily inaccessible to which entities. Are there any possible worlds at which Fido engages in contemplation? McMahan, Kagan, and Tooley would answer affirmatively, but others would disagree.<sup>38</sup> Are there any possible worlds at which the tree that is actually outside my window feels pleasure? Some would say ‘yes’, and others would say ‘no’.<sup>39</sup> If necessary inaccessibility excludes goodness, then we won't be able to determine which kinds are basic goods for which subjects without first settling difficult metaphysical issues like these. My account allows us to circumvent such issues precisely because it says that  $K$  can be a basic good for  $S$  even if  $K$  is necessarily inaccessible to  $S$ .

---

<sup>38</sup> See footnotes 12 and 13.

<sup>39</sup> Mills (2013, p. 23) claims that a tree could become sentient, and even rational.

My point is not merely that theorizing about welfare will be tidier and easier if we reject Necessary Inaccessibility Excludes Goodness—though it will be. It seems to me that whether there are any possible worlds in which Fido engages in contemplation is just irrelevant to whether it is a basic good for him. All that matters is whether, for any instances of contemplation, Fido would benefit directly from engaging in them if he were to engage in them. Recall the analogy to the theory of right action. We would not deny that there is a single, invariable list of wrong act types on the grounds that some agents are necessarily incapable of performing actions of some of those types. Why would necessary inaccessibility exclude goodness, if it doesn’t exclude wrongness?

Even if necessary inaccessibility *does* exclude goodness, however, it does not follow that variabilism is true. After all, it could be that none of the kinds that are basic goods for any welfare subjects are necessarily inaccessible to any other welfare subjects. This would be true, for example, if pleasure is the only basic good and every welfare subject is capable of pleasure. Absent further arguments (e.g., an argument that theoretical contemplation is a basic good for me but necessarily inaccessible to Fido), the truth of Necessary Inaccessibility Excludes Goodness would not disprove invariabilism.

But suppose we accept Necessary Inaccessibility Excludes Goodness and we also believe that some kinds that are basic goods for some welfare subjects are necessarily inaccessible to other welfare subjects. Then, in light of what I have argued, we should accept the following picture. There is a ‘master list’ consisting of the basic goods for the most advanced welfare subjects—those to whom no kind that is a basic good for any subject is necessarily inaccessible. The list of basic goods for any subject *S* is the subset (possibly improper) of the master list that is possibly accessible to *S*. For example, if the master list consists of pleasure, achievement, friendship, and rationality, and only the last of these is necessarily inaccessible to *S*, then the basic goods for *S* are pleasure, achievement, and friendship. The implications of this view, which I call *shallow variabilism*, are not significantly different from those of invariabilism.

First, consider the possibility that the correct theory of our welfare is an objective list theory on which some of the basic goods are inaccessible to certain welfare subjects or not well-suited to their actual natures. Shallow variabilism agrees that actual inaccessibility doesn’t exclude goodness, and that a basic good needn’t be well-suited to a subject’s actual nature. Thus, like invariabilism, it says

that a theory of this sort could be true of all welfare subjects. What if the correct theory of our welfare is an objective list theory on which some of the basic goods are *necessarily* inaccessible to some welfare subjects? Unlike invariabilism, shallow variabilism would deny that this theory is true of all subjects. But this wouldn't create serious problems for the objective list theory: the correct theory of the welfare of a primitive subject could, in principle, be recovered by taking the subset of the objective list that is possibly accessible to it—though it might be difficult to do this in practice given how hard it is to know which kinds are necessarily inaccessible to which subjects. We would be committed to pluralism about the true theories of welfare, but on shallow variabilism, this would be relatively innocuous.

Second, shallow variabilism allows us to rule out the view that the correct theory of our welfare is a sophisticated theory. As I argued earlier, since many simple subjects like newborns and dogs are positive in welfare, each of those simple subjects must have some psychologically simple kind (e.g., pleasure) as a basic good. But if the basic goods for any simple subject are a subset of the basic goods for us, we must also have some psychologically simple kind as a basic good. Sophisticated theories deny this, so they are false of us.

Shallow variabilism can be contrasted with *deep variabilism*, on which there is a kind  $K$  and subjects  $S$  and  $S^*$  such that although  $K$  is possibly accessible to both  $S$  and  $S^*$ ,  $K$  is a basic good for  $S$  but not for  $S^*$ . If deep variabilism were true, then a sophisticated theory could be true of us even though it is false of simple subjects, since a basic good for simple subjects could fail to be a basic good for us. Moreover, if we accepted an objective list theory of our welfare, we would still have to identify the correct theories of the welfare of other subjects, and this task would not be straightforward even in principle (as it would be on shallow variabilism). My arguments count against deep variabilism, even if shallow variabilism turns out to be true. Even if you accept Necessary Inaccessibility Excludes Goodness, you should reject deep variabilism because of its greater complexity and because its truth cannot plausibly be explained.

To summarize: we should reject Necessary Inaccessibility Excludes Goodness, both because it is intuitively false and because it would hold the philosophy of welfare hostage to metaphysical fortune. But even if we accept it, we can still be invariabilists if we maintain that no basic good for any welfare subject is necessarily inaccessible to any other welfare subject. Finally, even if we accept

Necessary Inaccessibility Excludes Goodness and we believe that some kind that is a basic good for some welfare subjects is necessarily inaccessible to others, we should accept shallow variabilism—a view whose implications are largely the same as those of invariabilism.

## 5. Conclusion

I have argued that we should accept invariabilism, and I have explained why we should therefore reject all sophisticated theories of welfare. If I am right, then the correct theory of welfare posits at least one basic good whose tokens can be possessed by the simplest welfare subjects. Provided that every welfare subject is capable of pleasure and desire, hedonism and desire satisfactionism are candidates for the true theory.<sup>40</sup> But many are persuaded by the familiar objections to those views, or convinced that some sophisticated kind is a basic good for us.<sup>41</sup> Others are inclined to believe that plants can be well off even though they are incapable of pleasure or desire. These philosophers should consider the following possibility: the true theory of welfare is an objective list theory on which there are a variety of basic goods requiring varying degrees of psychological sophistication, but at least one basic good that is accessible to the simplest welfare subjects. This is a promising view that merits further investigation.<sup>42</sup>

---

<sup>40</sup> If every welfare subject is also capable of belief, then subjective desire satisfactionism (Heathwood (2006)) and disjunctive desire satisfactionism (Lin (2016a)) are also viable candidates.

<sup>41</sup> See Lin (2014) and (2016c) for some objections to hedonism, and see Lin (2014), (2016a), and (2017) for some objections to desire satisfactionism.

<sup>42</sup> I thank Ben Bramble, Stephen Campbell, Richard Yetter Chappell, Nicolas Delon, Fred Feldman, Guy Fletcher, Johann Frick, Josh Gillon, Jennifer Hawkins, Chris Heathwood, Don Hubin, Antti Kauppinen, Charlie Kurth, Adam Lerner, Errol Lord, Barry Maguire, Kris McDaniel, Tristram McPherson, Duncan Purves, Jason Raibley, David Sobel, Valerie Tiberius, Travis Timmerman, participants at the 2015 Kansas Workshop on Well-Being and at the 2016 Syracuse Philosophy Annual Workshop and Network, audiences at Rice University and The Ohio State University, and the reviewers and editors at *Ethics*.

## Works Cited

Bradford, Gwen (2016), “Perfectionism,” in Guy Fletcher (ed.), *The Routledge Handbook of Philosophy of Well-Being* (Abingdon, UK: Routledge).

Bruckner, Donald (2016), “Quirky Desires and Well-Being,” *Journal of Ethics & Social Philosophy*, [www.jesp.org](http://www.jesp.org), vol. 10, no. 2.

Campbell, Stephen and Stramondo, Joseph (forthcoming), “Disability, Well-Being, and the Complicated Question of Neutrality,” *Kennedy Institute of Ethics Journal*.

Dorsey, Dale (forthcoming), “Why Should Welfare ‘Fit?’” *Philosophical Quarterly*.

\_\_\_\_\_ (2015), “The Significance of a Life’s Shape,” *Ethics* 125: 303-30.

\_\_\_\_\_ (2012), “Subjectivism without Desire,” *Philosophical Review* 121: 407-42.

Fletcher, Guy (2009), “Rejecting Well-Being Invariabilism,” *Philosophical Papers* 38: 21-34.

Heathwood, Chris (2006), “Desire Satisfactionism and Hedonism,” *Philosophical Studies* 128: 539-63.

\_\_\_\_\_ (2005), “The Problem of Defective Desires,” *Australasian Journal of Philosophy* 83: 487-504.

Kagan, Shelly (2016), “What’s Wrong with Speciesism?” *Journal of Applied Philosophy* 33: 1-21.

Kauppinen, Antti (2012), “Meaningfulness and Time,” *Philosophy and Phenomenological Research* 84: 345-77.

Kraut, Richard (2007), *What Is Good and Why: The Ethics of Well-Being* (Cambridge, MA: Harvard University Press).

\_\_\_\_\_ (1994), “Desire and the Human Good,” *Proceedings and Addresses of the American Philosophical Association* 68: 39-54.

Lin, Eden (unpublished), “Why Subjectivists about Welfare Needn’t Idealize”

\_\_\_\_\_ (forthcoming), “Enumeration and Explanation in Theories of Welfare,” *Analysis*.

Eden Lin – “Welfare Invariabilism” (forthcoming in *Ethics*)

\_\_\_\_\_ (2017), “Against Welfare Subjectivism,” *Noûs* 51: 354-77.

\_\_\_\_\_ (2016a), “The Subjective List Theory of Well-Being,” *Australasian Journal of Philosophy* 94: 99-114.

\_\_\_\_\_ (2016b), “Monism and Pluralism,” in Guy Fletcher (ed.), *The Routledge Handbook of Philosophy of Well-Being* (Abingdon, UK: Routledge).

\_\_\_\_\_ (2016c), “How to Use the Experience Machine,” *Utilitas* 28: 314-32.

\_\_\_\_\_ (2014), “Pluralism about Well-Being,” *Philosophical Perspectives* 28: 127-54.

McDaniel, Kris (unpublished), “The Metaphysics of Axiology and the Welfare of Animals.”

McMahan, Jeff (2002), *The Ethics of Killing: Problems at the Margins of Life* (Oxford: Oxford University Press).

Mills, Eugene (2013), “Early Abortion and Personal Ontology,” *Acta Analytica* 28: 19-30.

Nolan, Daniel (2013), “Impossible Worlds,” *Philosophy Compass* 8: 360-72.

Raibley, Jason (2013), “Values, Agency, and Welfare,” *Philosophical Topics* 41: 187-214.

\_\_\_\_\_ (2010), “Well-Being and the Priority of Values,” *Social Theory and Practice* 36: 593-620.

Rosati, Connie (2009), “Relational Good and the Multiplicity Problem,” *Philosophical Issues* 19: 205-34.

\_\_\_\_\_ (2006), “Personal Good,” in Terry Horgan and Mark Timmons (eds.), *Metaethics After Moore* (Oxford: Clarendon Press).

Scheffler, Samuel (2011), “Valuing,” in Jay Wallace, Rahul Kumar, and Samuel Freeman (eds.), *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon* (Oxford: Oxford University Press).

Sumner, L. W. (1996), *Welfare, Happiness, and Ethics* (Oxford: Clarendon Press).

Tiberius, Valerie (2014), “How Theories of Well-Being Can Help Us Help,” *Journal of Practical Ethics*

Eden Lin – “Welfare Invariabilism” (forthcoming in *Ethics*)

2(2): 1-19.

Tooley, Michael (1972), “Abortion and Infanticide,” *Philosophy & Public Affairs* 2: 37-65.

Velleman, J. David (1991), “Well-Being and Time,” *Pacific Philosophical Quarterly* 72: 48-77.

Yelle, Benjamin (2016), “In Defense of Sophisticated Theories of Welfare,” *Philosophia* 44: 1409-18.

\_\_\_\_\_(2014), “Alienation, Deprivation, and the Well-Being of Persons,” *Utilitas* 26: 367-84.