Resisting buck-passing accounts of prudential value

Guy Fletcher

Abstract This paper aims to cast doubt upon a certain way of analysing prudential value (or *good for*), namely in the manner of a 'buck-passing' analysis. It begins by explaining why we should be interested in analyses of *good for* and the nature of buck-passing analyses generally (§I). It moves on to considering and rejecting two sets of buck-passing analyses. The first are analyses that are likely to be suggested by those attracted to the idea of analysing *good for* in a buck-passing fashion (§II). The second are the buck-passing analyses of *good for* proposed by John Skorupski (§III), Henry Sidgwick (§IV), and Stephen Darwall (§V). Along the way the paper shows that Michael Smith's and Peter Railton's analyses of other concepts—analyses that could be (and have been) taken to be analyses of *good for*—are similarly unsuitable as analyses of it. The paper concludes by suggesting that the fact that none of the buck-passing accounts of *good for* considered here is satisfactory, coupled with an appreciation of the various problems that a buck-passing analysis of *good for* would have to avoid, suggests that we should be sceptical about the prospects of finding such an analysis and should look for one of a different type.

Keywords Buck-passing · Good · Good for · Prudential value · John Skorupski · Michael Smith · Peter Railton · Henry Sidgwick · Stephen Darwall

1 Introduction

We are all interested in what is good for ourselves, our friends and family, and people in general. This is reflected in the fact that *the* good for (that which is good

G. Fletcher (✉)
Exeter College, Oxford, UK
e-mail: Guy.Fletcher@Exeter.ox.ac.uk

for) is a part of almost every commonly-discussed moral theory.[1] Utilitarians tell us to maximise *the* good for, seeing this as the only basic moral requirement.[2] And almost all teleological theories hold that *the* good for is at least part of the good to be promoted. Although teleologists have placed more emphasis on *good for* in their theories, this is not to suggest that deontologists are not, or should not, be interested in it. However much they differ as to the weight to be given to such a duty, understanding what *good for* is, and which things are good for people will be necessary to give content to the duty of beneficence that any plausible deontological theory will have.

Despite the fact that *good for* is clearly important, and the fact that there has been a lot of work done on trying to determine which things are good for people, there has been comparatively little work on the metaethics of *good for*—of what it is for something to be good for someone. Philosophers often use *good for* without providing much explanation of how it is to be understood. Two examples are Roger Crisp and Fred Feldman who write:

> I wish to discuss Hedonism as a theory of well-being, that is, of what is ultimately *good for* any individual.[3]

> I want to know, in the abstract, what features make a life a good one *for the one* who lives it.[4]

Tim Scanlon provides a negative characterisation as follows:

> The intuitive notion of well-being that I am concerned with, then, is an idea of the quality of a life for the person who lives it that is broader than material and social conditions, at least potentially broader than experiential quality, different from worthiness or value, and narrower than choiceworthiness all things considered.[5]

However, some have attempted to go further than this by providing a positive analysis of *good for* in the form of a buck-passing analysis. And many of those who have not explicitly considered the issue are inclined, when pressed, to suggest similar analyses. For that reason, this paper examines both the buck-passing analyses that have actually been proposed and those that are likely to be suggested. The rest of this paper shows that no extant buck-passing analysis of *good for* is successful and that the kinds of buck-passing analysis that are likely to be offered are also inadequate.

---

[1] Adapting Moore's (1903/2000) terminology, I use '*the* good for' to refer to that which is good for.

[2] Writing about *good for* is replete with synonyms. At least as used by some writers, these include: prudential value, personal good, personal welfare, utility, well-being, and welfare.

[3] Crisp (2006, p. 100) (My italics).

[4] Feldman (2004, p. 13) (My italics).

[5] Scanlon (1998, p. 113).

2 Buck-passing accounts of value and prudential value

Buck-passing views are species of the 'fitting-attitude analysis' genus, one historically associated with Brentano and Ewing.[6] A buck-passing analysis of *good* claims that the property of goodness just is the property of having base properties that provide reasons for some pro-attitude, where this can be a mental state such as desiring, or an action such as promoting.[7] T. M. Scanlon formulated such a view thus:

> Goodness is not a single substantive property [...] rather to call something good is to claim that it has other properties (different ones in different cases) that provide such reasons.[8]

Buck-passers claim that the property of goodness is identical to the higher-order property of having properties that provide reasons to hold pro-attitudes towards the thing. They also claim that the goodness of something does not provide reasons to hold pro-attitudes towards it. Rather, it is the base or good-*making* properties that provide such reasons. Thus, the buck-passing account of goodness is a combination of two theses:

(1) The property of being good just is the property of having properties that provide reasons to adopt pro-attitudes towards the thing that is good.
(2) The property of being good does not, itself, provide a reason to adopt pro-attitudes towards the thing that is good.[9]

In short, the buck-passing view of *good* is:

> *Buck-Passing account of good* (1): T is good = T has base properties that give everyone reasons to adopt a pro-attitude towards T.

From this, we might formulate a first buck-passing analysis of *good for* thus:

> *Buck-Passing account of good for* (1): T is good for X = T has base properties that give everyone reasons to adopt a pro-attitude towards T.

The problem with this is that it does not distinguish between things that are good and things that are good for X. Thus we need something more like the following:

> *Buck-Passing account of good for* (2): T is good for X = T has base properties that give X reasons to adopt a pro-attitude towards T.

Is *Buck-Passing account of good for (2)* a plausible analysis of *good for*? No, for the simple reason that it claims that *anything* that X has a reason to adopt a pro-attitude towards is good for X. But X has reasons to adopt pro-attitudes towards many things that are not connected with what is good for X, such as the good for other people.

---

[6] Brentano (1889), Ewing (1947, 1959).

[7] The term 'pro-attitude' is credited to Ross by Ewing (1947, p. 149).

[8] Scanlon (1998, p. 11).

[9] Buck-passers often claim that (2) follows from (1). For discussion of this see Schroeder (2009).

Thus the analysis is insufficiently discriminating. To get around this, someone might suggest the following:

> *Buck-Passing account of good for* (3): T is good for X = T has base properties that give X <u>special</u> reasons to adopt a pro-attitude towards T.

The important question is how to understand *special* reasons. One suggestion is that the reasons provided by things that are good for X are of a distinct kind. For example, some might claim that what characterises the reasons provided for X by T is that they are prudential reasons, or reasons of self-interest.[10] There are two problems with this.

First, the analysis leaves the nature of *good for* almost as mysterious as before. If we are aim to understand what it is for something to be good for X, we learn little in being told that it is its giving X prudential or self-interested reasons. This is especially clear when we note that one of the ways in which people refer to *good for* is as *prudential value*. As such, this rendering of 'special' as prudential is not illuminating as to what it is to be *good for*. It simply invites us to ask of prudential reasons an analogous question to what we began by asking of *good for*: what is it for something to be a *prudential* reason?

Second, there is a problem with the order of explanation. If there is a distinct kind of reason—a prudential one—then what makes it so is that the consideration that favours the response is (or is related to) something good for the agent. As an analogy, take aesthetic reasons. What makes aesthetic reasons *aesthetic* is that the facts that favour certain responses are facts about *aesthetic* items (artworks etc.).[11] We go from the kind of consideration to the kind of reason generated. So just as we learn little about what it is for something to be an artwork by learning that it is to be something that generates aesthetic reasons, we make no progress in understanding what it is for something to be good for someone by learning that it generates prudential reasons.

Instead of the special reasons referred to in *Buck-Passing account of good for* (3) being of a distinct *kind*, could they be reasons with a special strength or priority for the agent? So what it is for something to be good for X is for it to provide X with reasons to adopt a pro-attitude towards it that are stronger than the reasons that X has to adopt a pro-attitude towards other good things. Put another way, someone's pleasure's being good for them just is their having reasons to adopt pro-attitudes towards it that are stronger than the reasons they have to adopt pro-attitudes towards similar pleasures of other agents (and vice versa).

The problem with this is that, even if we assume *arguendo* that agents have special reasons to (e.g.) promote their own pleasure, it is plausible to think they have at least equally strong reasons to promote the pleasures of *at least some* other people—family and friends for example. If so, *Buck-Passing account of good for* (3) yields the implausible result that the pleasures of their family and friends are good

---

[10] Such a suggestion is found in Parfit (forthcoming).

[11] Alternatively, what makes aesthetic reasons *aesthetic* (and epistemic reasons *epistemic*) is the kind of *responses* the properties of the thing favours, such as some kind of appreciation (or belief). The question, then, is whether there is some distinctive kind of response that characterises some reasons as prudential. It is hard to see that there is one.

for them. Furthermore, it is better to have an analysis of *good for* that leaves this kind of issue as open as possible, rather than one that takes a stand on whether we have special reason to favour certain others as much as, less than, or more than, ourselves. Thus trying to understand the special reasons that feature in *Buck-Passing account of good for* (3) in terms of the strength of the reason is not a promising strategy.

A third interpretation of the special reasons invoked in *Buck-Passing account of good for* (3) is that they are reasons for X to adopt a pro-attitude towards T that remain in place even if other agents do not have reasons to adopt a pro-attitude towards T. One consequence of this interpretation is that, when combined with a buck-passing account of good (in which for something to be good is for it to have properties that provide *everyone* with reasons to adopt a pro-attitude towards it), something that is good for someone need not be good. For example, a wicked person's prosperity is something that he has reasons to adopt a pro-attitude towards, and so is good for him according to *Buck-Passing account of good for* (3), even though, because others do not have reasons to adopt a pro-attitude towards it, it is not good according to the buck-passing analysis of *good*.[12]

One objection to understanding 'special' in counterfactual terms is that—like understanding special in terms of the *strength* of the reason—it is insufficiently discriminating. Whatever plausibility there is in the suggestion that an agent could have reason to promote their own pleasure even when it is not good, there is equal plausibility in the claim that they could have reason to promote the pleasure of at least some others, even when that pleasure is not good. For instance, an agent X might have special reasons of this counterfactual sort to promote the pleasure of Y, because of a relationship in which they stand. If this is so, and if we assume that not all such special reasons can plausibly be understood as relating to *the* good for X, understanding 'special' in this counterfactual manner does not work either.

We have seen that, combined with *any* of the foregoing interpretations of 'special', *Buck-Passing account of good for* (3) is unsatisfactory. Someone might try to avoid these problems by suggesting *Buck-Passing account of good for* (4):

> *Buck-Passing account of good for* (4): T is good for X = T is a mental state *of* X that has base properties that give X *special* reasons to adopt a pro-attitude towards T.[13]

Imagine, for example, that T is a state of pleasure of X. T is a mental state of X and T provides X with special reasons to adopt a pro-attitude towards it. One virtue of *Buck-Passing account of good for* (4) is that it is not vulnerable to the cases that derailed *Buck-Passing account of good for* (3). *Buck-Passing account of good for* (4) avoids these problem because X's having special reasons to promote the pleasure of X's children does not deliver the result that such pleasures are, themselves, good for X. This is because the pleasures of X's family and friends are obviously not themselves mental states of X.

---

[12] This example is taken from Dancy (2004, p. 173).

[13] This leaves open which account of *special* the analysis uses, if any.

A problem with this analysis is that it will settle the first-order question—namely, which things are good for people—prematurely. For example, it rules out the possibility of anything other than mental states being part of *the* good for. One might claim that this is just because the analysis was needlessly formulated too restrictively and that we can simply replace *Buck-Passing account of good for (4)* with the following:

> *Buck-Passing account of good for* (5): T is good for X = T is a physical or mental state of X, or […] that has base properties that give X *special* reasons to adopt a pro-attitude towards T.

This handles the previous objection in particular but only at the expense of generating other problems. First, *Buck-Passing account of good for (5)* is incomplete so it is not an analysis. At best it is a proposal for an analysis, one that would have to be filled out by expanding it to cover every plausible candidate for part of the good for.

The second problem, one that would be encountered if one tried to actually develop this analysis in that fashion, is that the more the analysis is expanded to cover plausible candidates for the good for the more likely it is to generate counter-examples.

A third problem is that such an analysis, even if it included every plausible candidate for the good for (and did not generate counter-examples), would not give an underlying unity to the things that can be good for. That is, we might hope for something that explains why mental and physical states alone (or whatever the full list contains) are the only kind of things that can be good for a person and why certain other kinds of things cannot be. So even if *Buck-Passing account of good for (5),* suitably expanded, got all and only the candidates for parts of *the* good for, this is a less illuminating analysis than one that explains *why* the things that plausibly can be part of *the* good for can be so. Thus, the problems with *Buck-Passing account of good for (5)* are: it is not an analysis of *good for*; it might not be capable of getting the right answers without generating counterexamples; and, even if it could, it would provide little explanation of the nature of *good for*.

A different tack someone might take is to appeal to an ideal observer, as in the following analysis:

> *Buck-Passing account of good for* (6): T is good for X = T has base properties, of the sort that a beneficent other would be sensitive to, that give X *special* reasons to adopt a pro-attitude towards T.

A problem here (one that also applies to the previous analysis) stems from the fact that it is possible for an agent to have special reasons to promote *the* good for at least some others to whom they stand in a special relationship. Assume for simplicity that pleasure is the only thing that is good for Edward (and likewise Joan). If a beneficent other would thus be attuned to pleasure then this, coupled with Joan's having *special* reasons to promote the pleasure of Edward, means that *Buck-Passing account of good for (6)* delivers the result that the pleasure of Edward is good for Joan. *Buck-Passing account of good for (6)*, like earlier analyses, does not home in on *the* good for Joan sufficiently. To avoid this problem, someone might

suggest identifying what it is to be good for someone with being something that they would have reasons to promote even if they were the only person:

> *Buck-Passing account of good for* (7): T is good for X = T has base properties (that exist even if only X exists) that give X *special* reasons to adopt a pro-attitude towards T.

One problem for *Buck-Passing account of good for* (7) is that it rules out certain plausible candidates for *the* good for. One example is that friendship and the possibility of any other-involving good, such as romantic love, being part of *the* good for is incompatible with *Buck-Passing account of good for* (7). If friendship can be good for X then we have a violation of the left-to-right reading of the identity statement. This is because friendship would be good for X but there will not be base properties associated with friendship in a world in which only X exists. Even if everything that generates reasons for X in a world in which only X exists is good for X, the reverse does not hold. Thus we should reject *Buck-Passing account of good for* (7) because it rules out plausible candidates for the *good for*.

Having examined, and rejected, some buck-passing analyses of *good for* that are likely to be suggested, it is necessary now to examine, and see the grounds for rejecting, some of the analyses of *good for* that have been proposed, namely those of John Skorupski, Henry Sidgwick, and Stephen Darwall.

3 Skorupski (and Smith and Railton)

John Skorupski presents an analysis of *good for* as follows:

> In general, I would argue, the good of a person comprises what there is sufficient (evaluative) reason for that person to desire, or what there would be such reason for them to desire if they were capable of sufficiently developed desires. […] In this case, then, the pro-act is desiring, and the reason for favouring is agent-relative. Hence also the good in this case is also agent-relative: we talk of the good of, or for, x, and by that we mean whatever there is reason for that person, x, to desire.[14]

Skorupski's view is that being good for someone is to be something that they have agent-*relative* reasons to desire. If we understand the 'specialness' of the reasons as their being agent-relative, then Skorupski's analysis is an instance of *Buck-Passing account of good for* (3):

> *Buck-Passing account of good for* (3): T is good for X = T has base properties that give X special reasons to adopt a pro-attitude towards T.

And Skorupski's view fails on the same ground as that analysis because it is too inclusive. Agents have agent-relative reasons to adopt pro-attitudes towards things that are not good for them (such as the good for their friends and family). So Skorupski's analysis is too broad. Even if it provides a necessary condition for

---

something's being good for X, it does not give us necessary and sufficient conditions.

It is useful to note that the same objection, that of being too broad, would apply to Michael Smith's account of agent-relative good if this were treated as an account of *good for*.[15] Smith claims that goodness is a relational property—which he represents as good$_x$—and that something's being good$_x$ is its being something X would desire if X had a set of desires that was maximally informed, coherent, and unified. But whatever plausibility this has for the general category of agent-relative value, it is too broad as an analysis of *good for*. As Smith writes:

> Note that I have claimed here that being F is good for A only if A's being F is good$_A$, *not* that being F is good for A *if and only if* A's being F is good$_A$. […] A's being F might well be good$_A$ without in any way contributing to A's welfare.[16]

The same holds for Peter Railton's account of *an individual's good*. Railton's view is that X's good consists in what a fully-informed and instrumentally rational counterpart of X (who contemplated X's situation) would want X to want. Some have interpreted Railton's analysis as an account of *good for* but he makes clear why it cannot plausibly be regarded as one:

> This notion is not the same as that of an individual's welfare, for it may turn out that an ideally informed and rational individual would want to seek as an end in itself (were he to step into the place of his present self) the well-being of others as well as himself.[17]

We cannot construe *good for* as that which an agent has reason to desire nor, as Railton and Smith concede, as what a fully-informed counterpart of the agent *would* desire, for that encompasses things that are not parts of the good for that agent, such as the good for others. Thus Railton's and Smith's analyses would not work as analyses of *good for*.[18] And Skorupski's buck-passing analysis of *good for* is inadequate on the same ground: it is too broad.

4 Sidgwick's account

Henry Sidgwick provides an account of *good for* as follows:

> [W]e may keep the notion of 'dictate' or 'imperative' merely implicit and latent, – as it seems to be in ordinary judgements as to 'my good' and its opposite – by interpreting 'ultimate good on the whole for me' to mean *what I*

---

[15] Smith (2003).

[16] Smith (2003, p. 584) (My italics).

[17] Railton (2003, p. 54 n. 9). C.f. Rosati (1996, p. 298 n.2).

[18] I should point out that in addition to not being analyses of *good for*, Railton's and Smith's analyses are arguably not buck-passing analyses. I include them because they could be taken to be analyses of *good for* (and Railton's sometimes is) and because they are useful in demonstrating the breadth problem for Skorupski's account.

should practically desire if my desires were in harmony with reason, assuming my own existence alone to be considered.[19]

This can be interpreted as a buck-passing account that identifies something being good for X with being something X has reason to desire 'assuming his own existence alone to be considered'. A vital question for assessing Sidgwick's proposal is how to understand 'my own existence alone to be considered'. One way is as a kind of isolation test, motivated by the thought that because the presence of agents other than X makes it hard to narrow in on good for X in analysis we should imagine that there are no other agents and see what reasons remain. We can interpret Sidgwick's analysis with this interpretation thus:

Sidgwick (1): T is good for X = T has base properties (that would exist if only X exists) that give X reasons to adopt a pro-attitude towards T.

This is similar to Buck-Passing account of good for (7):

Buck-Passing account of good for (7): T is good for X = T has base properties (that exist if only X exists) that give X special reasons to adopt a pro-attitude towards T.

Sidgwick (1) fails as an analysis for similar reasons as *Buck-Passing account of good for (7)*. First, it is not clear that it would only generate reasons concerning *the* good for X. Even in a world in which X is the only agent, X can have reasons that do not concern the good for X, such as reasons with respect to the natural environment.

Second, it might also be the case that *Sidgwick (1)* rules out plausible candidates for things that are good for X, such as friendship. In worlds in which X is the only agent, X cannot have reasons to (e.g.) promote friendship—on the assumption that friendship is impossible in such a world. Thus *Sidgwick (1)* rules out the possibility of, not only friendship, but also love, and any other interpersonal kinds of thing being good for X. But given that at least some such interpersonal relations are good for people, *Sidgwick (1)* cannot be correct.

In objection, someone may suggest that X has reasons to *desire* friendship in this scenario, for instance, and that this allows Sidgwick's analysis to avoid the second objection. This will not work, however, because if we allow reasons to desire into the analysis there is nothing to stop someone from having reasons to desire that any other people be happy. Or, even if they know that they are the only living person, they would still have reasons to desire that other people exist and be happy without that being good for them. Thus, the second objection to *Sidgwick (1)* still holds. We should reject an isolation test form of the analysis.

Another interpretation of Sidgwick's analysis is suggested by his writing:

To simplify the discussion, we will consider only what a man desires for itself – not as a means to an ulterior result, – *and for himself – not benevolently for others: his own Good and ultimate Good.*[20]

---

[19] Sidgwick (1907/1981, p. 112) (My italics).

[20] Sidgwick (1907/1981, p. 109) (My italics).

Sidgwick perhaps takes 'assuming my own existence alone to be considered' to be equivalent to 'for himself—not benevolently for others'. The question now is how to understand a person's desiring something 'for himself, not benevolently for others'. Presumably, Sidgwick thinks that for an agent to desire something for himself, as opposed to benevolently for others, is simply to desire that *he* have the thing in question. We can re-render Sidgwick's analysis thus:

> *Sidgwick (2)*: S is good for X = S has base properties that give X reasons to adopt a pro-attitude towards his having S.

The problem with this is that there are many instances of things that are good or bad for people that do not satisfy this condition. In the case of someone who suffers from a mental disability one aspect of what is bad about their condition is precisely the fact that it makes them *unable* to adopt the kinds of attitudes that *Sidgwick (2)* includes. For example, they will be unable to adopt a pro-attitude towards some improvement in their condition or con-attitudes towards some element of their condition or to a deterioration in their condition. One might try to counter this by insisting that for any putative example that either:

> (a) The person *can* adopt the relevant pro/con-attitudes to S.
> Or
> (b) S is not good/bad for X.

But this manoeuvre is incapable of handling every putative case. For example, it does not seem plausible in the case where part of what is bad about a mentally disabled person's condition is their being unable to adopt certain kinds of attitude. As such Sidgwick (2) fails as an analysis of *good for*. And in the absence of another plausible way of interpreting Sidgwick's analysis, we must conclude that it fails.

5 Darwall's Rational-Care Theory

The final buck-passing account of *good for* considered here is Stephen Darwall's 'Rational-Care Theory of Welfare' (hereafter *Rational Care Theory*).[21] Darwall presents it thus:

> We might equivalently say that what it is for something to be good for someone is for it to be something that is rational (makes sense, is warranted or justified) to desire for him insofar as one cares about him.[22]

To understand it we need to know how to understand the 'insofar as one cares about him' clause.[23] One way to interpret it is as the claim that something is good for X if

---

[21] Darwall (2004). Darwall uses 'welfare' but (quotations apart) I retain my terminology for the sake of continuity. Darwall (2006a, p. 434) explicitly claims that the terms label the same concept.

[22] Darwall (2004, p. 9).

[23] For discussion of issues with the formulation of Darwall's theory see: Feldman (2006) and Hurka (2006).

we *would* have reason to want it for X if (and only if) we cared about him. Darwall denies this however:

> Insofar as we care for someone, we ought to be guided by his good. So far, these reasons are merely hypothetical. The idea, however, is not that the fact that one cares about someone makes considerations of his good reasons for one. *The reasons are not conditional on one's caring. If that were so, they would be cancelled once one ceased to care.*[24]

He then adds the crucial claim for understanding his analysis. Concerning the reasons to be guided by a person's good:

> They are conditional, rather, on a hypothesis one accepts or is committed to *in* caring, namely, that the cared for is *worth* caring for.[25]

Darwall's claim is that a necessary part of caring is the belief that the person is *worth* caring for. It is that Y is *worth* caring for—the hypothesis X accepts in caring for Y—that makes the good for Y reason-providing. Thus, that which is good for Y makes a claim on Z, even if Z does not in fact care for Y, because Y is *worth* caring for. Darwall makes this clear thus:

> a person has reason to care about his own good because he has reason to care for himself. And he has reason to care for himself because he, like any person, has worth – he matters.

Having clarified Darwall's view, we need to test its adequacy. Starting with Darwall's statement:

> We might equivalently say that what it is for something to be good for someone is for it to be something that is rational (makes sense, is warranted or justified) to desire for him insofar as one cares about him.[26]

We can formulate the view thus:

> *Darwall (1):* Y is good for X = Y has base properties that give someone who cares for X reasons to desire X's having Y.

Unfortunately, the analysis faces counterexamples. It cannot ever allow the combination of X's caring for Y and X's having reason to desire that Y have something that is not good for Y. Yet people can have such reasons. For example, we all have reasons to want people that we care for (indeed, everyone), to have certain virtues and to act rightly. And it is clear that our doing so does not depend upon its being good for them to have these virtues or to act rightly. This is just one instance of the general problem that the analysis rules out a plausible class of reasons that people who care about others have that do not relate to what is good for that person.[27]

---

[24] Darwall (2004, p. 8) (My italics).

[25] Ibid (Italics in original).

[26] Darwall (2004, p. 9).

[27] For further discussion see Raz (2006, p. 413).

Some may dig their heels in against such counterexamples so it is necessary to point out a second and more significant problem. This is that the analysis is problematically circular. So *even if* Darwall's analysans has the right extension it is not a particularly illuminating account of *good for*. Anticipating such a charge, Darwall admits that 'care cannot be defined in terms of welfare, on pain of circularity' and asks 'Don't we need to understand the idea of welfare first before we can understand what care is?'[28] To see the problem, start with the earlier formulation:

> *Darwall (1):* Y is good for X = Y has base properties that give someone who cares for X reasons to desire X's having Y.

This will be in trouble if caring for X *just is* desiring the good for X because good for X will then appear on the right-hand side of the analysis thus:

> *Darwall (1*):* Y is good for X = Y has base properties that give someone who desires the good for X reasons to desire X's having Y.

Even if the analysans has the right extension, *Darwall (1*)* tells us little about what it is for something to be good for someone, for it is unilluminating to be told that what it is for something to be good for someone is for it to be something that someone who desires what is good for someone would have reason to desire that they have. We learn that there is a connection between *good for* and *care* but little else, and certainly less than we would want from an analysis of *good for*. To stave off this objection Darwall needs to show that *Darwall (1)* is not equivalent to *Darwall (1*)* by showing that care is not equivalent to desiring a person's good.

Darwall's first response to this objection is that caring for someone is not *merely* desiring their good, for we can desire a person's good 'through whim or fancy, without involving any concern for the person himself' and that care:

> involves concern for another in light of apparent obstacles to her welfare. We desire her good, not just intrinsically (for *its* sake), but also for *hers*.[29]

Darwall also claims that care involves a whole host of emotional dispositions 'that need not accompany any intrinsic desire, even a desire for someone's benefit'.[30] This does not solve the problem, however, because it only refutes the claim that the analysis is circular because the concept of care is *identical to* the concept of desiring the good for the person. But the objection can be reformulated as the claim that a necessary part of the concept of care is the concept of desire for the good for the person cared for. And this is clearly true, given the unintelligibility of someone claiming that they cared about someone but wished only that they suffered, or had no desires for the good for them at all. And as long as a desire for the good for the person cared for is part of the concept of *care*, the circularity problem remains.

Darwall's second response to the problem of circularity is to suggest that care is (something like) a natural kind:

---

[28] Darwall (2004, pp. 47, 12).

[29] Darwall (2004, p. 69) (Italics in original).

[30] Darwall (2006b, p. 651).

> However, we need not define care…if it is something like a psychological natural kind. Just as we can use a term like 'water' without a prior definition to refer to the natural stuff in the rivers and lakes for empirical inquiry, so likewise might we refer to care for purposes of a Metaethical theory of welfare if it is a natural kind.[31]

For this strategy to work we need to be able to *identify* care without using the concept *good for*. Darwall attempts this task but does not succeed in distinguishing care from similar psychological states, such as empathy, except by invoking the concept of *good for*. He himself makes clear how difficult it is to distinguish care from empathy. This is because both attitudes have another person and their plight as their object and involve imaginative appreciation of their plight. There is this difference:

> Someone in the grip of resentment, envy or the desire for revenge may take delight in the vivid appreciation of another's plight he gets from imagining what another's situation must be like for her.

But in light of this, it looks like the only way to distinguish this kind of empathy from care is that in caring for someone one performs this act of imagining the person's situation from a concern for their good.[32] Given the similarities between care and empathy, which Darwall points out, we cannot individuate care except by reference to a desire for the good for the person cared for. Such a problem is unsurprising given that the desire for the good for the person cared for is an essential part of care (and there does not seem to be much more to care than the desire). Thus Darwall's suggestion of treating care as a natural kind does not work because we are unable to individuate care except by reference to *good for*. Thus the objection of being unilluminatingly circular stands and Darwall's analysis of *good for* fails.

It is worth briefly mentioning a different objection to Darwall's analysis. One might argue that even if the property Darwall's analysis picks out is necessarily coextensive with the property of being *good for*, these properties are non-identical because the former property is distinct and *results from* the property of *good for*. Thus, according to the objection, Y is good for X because of various base properties that Y has but that someone has reason to want X to have Y (if one cares about X) *because* it is good for X.[33] That is, the Darwall property—the property of being such that you have reason to desire that X have Y (if you care about X)—results from the distinct property of being good for X. As such the two properties are non-identical. This can be represented thus:

> Y is something that I have reason to desire that X have if I care about him
>
> | (Results from)
>
> Y is good for X
>
> |(Results from)
>
> Y has base properties (a, b, c…)

[31] Darwall (2004, p. 50).

[32] For further discussion of this kind of response see Shah (2004, p. 580).

[33] Kraut (2007, p. 182 n.32) mentions and endorses this objection to Darwall.

Though this claim has some plausibility it does rest upon explicitly rejecting a part of the buck-passing view of *good for* (the claim that *good for* is not reason-providing) in a way that the proponent of such a view might claim begs the question against them. Thankfully, we can leave the objection aside because the case against Darwall still goes through even if we allow that something is not reason-providing for those who care about a person *because* it is good for that person (but, instead, because of its base properties, such as being pleasurable). If we grant this claim to Darwall, then either:

> (a) We have two distinct properties with a common resultance base (the property of being good for X and the property of being something that one has reasons to desire for X, insofar as one cares about X)
> OR
> (b) We have the same property picked out in two different ways.

Alternative (a) would mean that Darwall's analysis fails to pick out the correct property—it picks out a property that is distinct from the property of being *good for*. Only if alternative (b) holds does Darwall's account pick out the correct property. My conclusion is that (a) might be true but that *even if* (b) is true, the analysis is still problematically circular and uninformative. For we cannot identify part of the analysans (*care*) except by invoking the notion that the analysis is meant to elucidate (*good for*). Even if the analysis picked out the property of *good for*, it only does so in a way that tells us little or nothing about it, so Darwall's buck-passing analysis of *good for* is inadequate.

6 Conclusion

In this paper we have seen various attempts to analyse *good for*—both suggested and actual—and various grounds on which they are inadequate. Someone might respond by pointing out that even if this shows that none of *these* analyses succeeds, some other buck-passing analysis might do so. But the failure of these analyses should lead us to question whether any such an analysis can avoid all of the problems that those covered here have faced. For a buck-passing analysis of *good for* to be successful it must: not prejudge substantive issues such as what reasons people have, and be compatible with plausible views as to the reasons that they have (such as the fact that agents have reasons to adopt pro-attitudes towards the good of at least some others). It must also not prejudge the substantive questions of which things are good for people, and be compatible with the fact that at least some of the things that are good and bad for people are things that prevent them from adopting attitudes. And it should be illuminating as to what it is for something to be good for someone and not circular. Seeing the ways in which the buck-passing analyses of good for considered here have failed to avoid these problems should lead us to be sceptical about the prospects for such an analysis. It should also generate motivation to investigate other types of analysis of *good for*.

References

Brentano, F. (1889). The origin of our knowledge of right and wrong.

Crisp, R. (2006). *Reasons and the good*. Oxford: Clarendon Press.

Dancy, J. (2004). *Ethics without principles*. Oxford: OUP.

Darwall, S. (2004). *Welfare and rational care*. Princeton: PUP.

Darwall, S. (2006a). Reply to Griffin, Raz, and Wolf. *Utilitas, 18*, 434–444.

Darwall, S. (2006b). Reply to Feldman, Hurka, and Rosati. *Philosophical Studies, 130*, 637–658.

Ewing, A. C. (1947). *The definition of good*. London: Routledge.

Ewing, A. C. (1959). *Second thoughts in moral philosophy*. London: Routledge and Kegan Paul.

Feldman, F. (2004). *Pleasure and the good life: Concerning the nature, varieties and plausibility of hedonism*. Oxford: OUP.

Feldman, F. (2006). What is the rational care theory of welfare? *Philosophical Studies, 130*, 585–601.

Hurka, T. (2006). A Kantian theory of welfare? *Philosophical Studies, 130*, 603–617.

Kraut, R. (2007). *What is good and why: The ethics of well-being*. Cambridge, Mass: HUP.

Moore, G. E. (1903/2000). *Principia ethica* (Revised ed.). T. Baldwin (Ed.), Cambridge: CUP.

Parfit, D. (forthcoming). *On what matters.* Oxford: OUP.

Railton, P. (2003). Facts and values. In *Facts, values and norms* (pp. 43–68). Cambridge: CUP.

Raz, J. (2006). Darwall on rational care. *Utilitas, 18*, 400–414.

Rosati, C. (1996). Internalism and the good for a person. *Ethics, 106*, 297–326.

Scanlon, T. M. (1998). *What we owe to each other*. Mass: HUP.

Schroeder, M. (2009). Buck-passers' negative thesis. *Philosophical Explorations, 12*(3), 341–347.

Shah, N. (2004). Review of Stephen Darwall, *Welfare and rational care*. *The Philosophical Review, 113*, 577–582.

Sidgwick, H. (1907/1981). *The methods of ethics* (7th ed.). Indiana: Hackett.

Skorupski, J. (2007). Buck-passing about goodness. In T. Rønnow-Rasmussen, B. Peterson, J. Josefsson, D. Egonsson (Eds.), [online resource]: *Hommage á Wlodek: Philosophical papers dedicated to Wlodek Rabinowicz*. Lund. http://www.fil.lu.se/HommageaWlodek/site/abstra.html.

Smith, M. (2003). Neutral and relative value after Moore. *Ethics, 113*, 576–598.