

**BEYOND WRONG REASONS:
THE BUCK-PASSING ACCOUNT OF VALUE**

Ulrike Heuer

University of Leeds

The buck-passing account of value (BPA) is very fertile ground that has given rise to a number of interpretations and controversies. It has originally been proposed by T.M. Scanlon as an analysis of value: according to it, being good 'is not a property that itself provides a reason to respond to a thing in certain ways. Rather, to be good or valuable is to have other properties that constitute such reasons'.¹ And also: 'being valuable is not a property that provides us with reasons. Rather, to call something valuable is to say that it has other properties that provide reasons for behaving in certain ways with regard to it.'²

As Pekka Väyrynen,³ and subsequently Mark Schroeder⁴ and Roger Crisp⁵ have pointed out, the account comprises two theses:

(BPA-) The fact that something is good or of value is not itself a reason to respond to it favorably or to behave in certain ways with regard to it.

¹ Scanlon 1998, p. 97

² Ibid, p. 96

³ Väyrynen (2006)

⁴ Schroeder (2009)

⁵ Crisp (2008)

(BPA+) The fact that something is good or of value consists in the fact that it has some other property P which is a reason to respond to it favorably or to behave in certain ways with regard to it.⁶ | ⁷

The two claims are in principle independent. As Crisp emphasizes (BPA-) does not entail (BPA+). Perhaps goodness is not a reason, but it does not follow from this that goodness must consist in there being other properties that provide reasons.⁸ The converse is also true: (BPA+) can be true, even if (BPA-) is false. Even if goodness consists in there being some other property P which provides a reason, goodness itself could also be a reason – in the way in which, if R is a reason, the fact that R obtains could also be a reason.⁹

Scanlon's original arguments attempt to show that (BPA+) is true because (BPA-) is. But these arguments are flawed.¹⁰ I will not be concerned with them here. The most influential of these arguments – the one that Crisp¹¹ aptly dubbed 'the redundancy argument' – has been extensively discussed, and shown to be unsuccessful.¹²

Of course the buck-passing account may be correct even if the arguments for it are not. But since (BPA-) and (BPA+) are independent of one

⁶ Perhaps the thesis should allow that goodness consists in a plurality or combination of other properties and thus BPA+ should read: 'The fact that something is good or of value consists in the fact that it has some other properties P1, P2, ... Pn which are reasons to respond to it favorably or to behave in certain ways with regard to it.'

⁷ The formulations are mine, but I take it that they are in keeping with Väyrynen's, Crisp's and Schroeder's views. Väyrynen, in his formulation of BPA+, emphasizes that goodness is supposed to be the higher-order property of other properties that provide reasons. I will discuss the property identity claim that this formulation brings out in section III.

⁸ Crisp wishes to argue that while BPA- is a plausible claim, we should not commit ourselves to BPA+. I disagree with him on this point in more than one way: I don't find BPA- plausible at all, but I also don't believe that there is much interest in it independent of its connection with BPA+.

⁹ For a more detailed argument, showing how this could be so, see my 'Wrongness and Reasons', as well as Schroeder (2009).

¹⁰ I have argued for this point in detail in Heuer (2006) and Heuer (2010), and so have Crisp (2005), Väyrynen (2006) and Schroeder (2009).

¹¹ Crisp (2005).

¹² In Heuer (2006), I called the argument 'the argument from explanation', but I think that Crisp's label is better, and more to the point, and I will therefore adopt it here.

another, there is a question about each of them why we should accept them. Some¹³ have focused on (BPA-) as the claim the buck-passing account is ultimately concerned with. Viewed thus, (BPA+) may have been introduced only as an explanation of (BPA-), and an unsuccessful one at that. I am inclined, however, to think that (BPA+) is the more central one of the two theses, and the theoretically more important one, as it suggests a metaphysical account of the relation of values and reasons. In this paper, I will therefore focus on (BPA+). I should add that while I am going to discuss (BPA+) as suggesting a metaphysical account, there is also an interpretation of buck-passing as a conceptual analysis of 'good', and of course the metaphysical and conceptual interpretations are compatible and could both be true.¹⁴ Mainly for simplicity's sake I will not run the two interpretations in tandem, but focus on the metaphysical one. (But I will draw your attention to the semantic alternative occasionally.)

In section I, I will show that BPA is not as obviously a successor of the fitting-attitude analysis (for short: FA analysis) of value as some have thought. The much discussed wrong-kind-of-reasons (for short: WKR) problem afflicts buck-passing only in so far as it incorporates a version of FA analysis, or at any rate is expressed in terms of reasons for attitudes. There can be a buck-passing account of value which is not affected by the problem: one that limits the account to reasons for actions.¹⁵ However, insofar as BPA does inherit elements of FA analysis, it also has a WKR problem. In section II, I will discuss this problem and its solution. I will show that it has been misidentified in the current literature, and that – once we understand the

¹³ As for instance Crisp (2008), also Parfit, if not quite as explicitly.

¹⁴ For instance, Rabinowicz and Rønnow-Rasmussen, as well as Daniellson and Olson take this view. Scanlon seems to vacillate between the two interpretations.

¹⁵ I take it that Parfit's version of BPA is an example of this.

problem correctly – its solution is likely to be unavailable to the buck-passer. Hence we should reject any account of BPA that incorporates FA analysis. That leaves us with versions which do not: versions that formulate BPA+ in terms of reasons for actions only, rather than reasons for attitudes. Finally, in section III, I will discuss at least briefly why buck-passing seemed to be appealing to begin with, and whether a version of BPA that does not incorporate FA analysis is a viable contender of the account – beyond the WKR problem.

I. Some background: Buck-passing and FA analysis

Buck-passing stands in a complicated relation to the fitting-attitude analysis of normativity that reaches back to the work of Brentano and Ewing¹⁶. Proponents of FA-analysis hold that, as a matter of conceptual analysis of ‘good’, something is good if it is fitting to have pro-attitudes of a certain kind towards it. The ambition of FA-analysis is to explain all practical normativity in terms of the fittingness of attitudes, which – while being itself a normative concept – is the only normative concept which grounds all others.

Wlodek Rabinowicz and Toni Rønnow-Rasmussen¹⁷ welcomed BPA as a contemporary version of FA-analysis, and others have followed this suggestion¹⁸. But whether BPA should be regarded as a version of FA-analysis is a moot point. First of all the second disjunct in both formulations above - the phrase ‘reasons... to behave in certain ways with regard to [something that is of value]’ - has no obvious role in FA-analysis. Let me begin therefore by dropping it for the moment and by focusing on the following shortened

¹⁶ Brentano (1934), Ewing (1947). – For a recent discussion see Bykvist (2009).

¹⁷ Rabinowicz / Rønnow-Rasmussen (2004).

¹⁸ See for instance, Daniellson and Olson (2007).

formulations of (BPA-) and (BPA+) (I will come back to the left out disjunct later):

(BPA_{FA} -) The fact that something is valuable is not itself a reason to respond to it favorably.

(BPA_{FA} +) The fact that X is valuable consists in the fact that X has some other property P that is a reason to respond to it favorably.

The FA-subscript indicates that in this form the formulations can be understood as part and parcel of FA-analysis. Or can they? It depends on whether the fact that there is a reason to respond favorably to something is the same fact as the fact that it is fitting to have a favoring attitude to that thing. Thus the question is whether (BPA_{FA}+) is equivalent to:

(FA_{BPA}) The fact that X is valuable consists in the fact that X has some other property P that makes it fitting to have certain favoring attitudes towards X.¹⁹

There are at least two worries about this:

First, not all reasons that we have in virtue of something being good are reasons to have favoring attitudes. Some of them are reasons to act. This is not a problem for BPA, because the second disjunct that I left out in the FA-formulations of the view are precisely about reasons for actions: in Scanlon's formulation, there are reasons 'to behave in certain ways towards something that is good'. Adding this expressly, as in (BPA-) and (BPA+), shows that the

¹⁹ I use the subscript 'BPA' for the version of FA, because it emphasizes that it properties *other than being valuable* that make the attitude fitting.

account is not restricted to reasons for attitudes. It also shows that there is no assumption that reasons for actions can simply be derived from reasons for attitudes. But FA-analysis would claim just that. If BPA+ were a version of FA-analysis, reasons for actions would have to be derived from reasons for attitudes. That the original BPA formulations distinguish between reasons for attitudes and reasons for actions may be owed to its author's skepticism about the FA-project²⁰. But if so, BPA lacks the neatness, and perhaps also part of the motivation that drives FA-analysis.

Secondly, there is an obvious difference between 'there is a reason to have attitude A' and 'it is fitting to have attitude A'. Reasons are, *pro tanto*, requiring:²¹ If there is a reason to have an attitude (and the reason is not defeated) I am at fault if I don't have it. Thus, if I have a reason to admire something (say) then, other things being equal, I am at fault if I fail to admire it. Fittingness is different: If I have an attitude, and it is fitting to have it, I am justified in having it, but it does not follow that I would have been at fault if I would not have had the attitude at all. And example might help to show this. Take fear: it is fitting to feel fear only if there is danger – but a person who is not afraid, even though she is aware of danger, is not always (if ever) at fault. In other words: That it is fitting to have an attitude is not a reason to have it. Thus, it is not obvious that FA-analysis could explain why a person may be required to have an attitude.²² The normative force of 'there is a reason to

²⁰ In his first book (1998), Scanlon claims that all reasons are reasons for attitudes – beliefs and intentions, respectively. But he does not defend this approach anymore. Scanlon (2008) claims that intentions are based on reasons for action, rather than the other way around.

²¹ Unless there are, as Jonathan Dancy (2004) has it, 'enticing reasons' – reasons which do not require compliance (other things being equal). However, if Dancy is right and there are enticing reasons in this sense, they wouldn't be relevant in our context, because BPA is concerned with all reasons for pro-attitudes, whereas 'enticing reasons' would be a subclass of those, distinguished by their content.

²² Joseph Raz suggested to me that a proponent of FA-analysis may claim that an attitude is required if it is fitting to have it and unfitting not to have it. This requires that 'unfitting' is

have an attitude' and 'it is fitting to have an attitude' appears to be different. Perhaps 'fittingness' is a normative concept, but if so, it is not clear how it translates into the language of reasons. It is not clear therefore that BPA is a version of FA-analysis at all, even in its BPA_{FA} formulations. None of this is a skin off the buck-passer's nose of course. There may be good reasons for departing from FA-analysis in the way she does. But those reasons have yet to be uncovered.

But whether or not buck-passing is a version of or a development from FA-analysis, it inherits one of its problems. The so-called wrong kind of reasons (for short: WKR) problem.

II. The WKR problem

1. The problem

The so-called WKR problem is the problem that, intuitively, there seem to be reasons to have favoring attitudes towards objects which are devoid of value as, for instance, when an evil demon orders you to admire him or else he'll torture you. You then have a reason to admire him, but he is not in any way good. Hence, the buck-passing account is false. The buck-passer's solution to the problem must be to narrow down the kinds of reasons of which BPA is true. Put schematically, the solution to the problem is to reformulate (BPA+) in a WKR-proof way, such as:

different from the negation of fitting (different from 'non-fitting'). Or rather: if it is fitting to have an attitude, it is not *eo ipso* unfitting not to have it – as the example of fear above has shown. It may be fitting to be afraid, but it is not unfitting not to be afraid. In those cases, where have the attitude is fitting and it is also the case that it would be unfitting not to have it, the agent may have a reason to have the attitude. - If the proponent of FA-analysis can make sense of these concepts and their logical relations, she may have a way of analysing 'required' – but as this shows, the explanation is not straightforward.

The fact that X is valuable consists in the fact that X has some other property P which provides a reason *of the right kind* to respond to it favorably.

The difficulty is to describe the relevant kind of reasons in a non-circular way. Obviously, saying that only those reasons that obtain when (and because) the attitude's object is of value will not do. In the semantic interpretation, this would make BPA circular; in the metaphysical one, the account would become uninformative and uninteresting.²³

There have been a number of attempts to solve this problem. Early on, Wlodek Rabinowicz and Toni Rønnow-Rasmussen suggested that a reason is of the right kind if it has the 'dual role' of providing both the content of the pro-attitude and of justifying it: E.g. that she is courageous is a reason for admiring her, and she is admired for her courage.²⁴ The evil demon on the other hand is admired on account of his threat (the avoidance of the threatened punishment is the justifying reason for having the admiring attitude), but he is not admired *for* being threatening (in this case, it seems quite unclear what, if anything, he is being admired for). However, Rabinowicz and Rønnow-Rasmussen themselves are sceptical about the solution, because it gets in trouble if the demon requires us to admire him for his own sake "on account of his determination to punish us if we don't". In this case, "the demon's determination to punish us if we don't comply provides the reason for our admiration and at the same time appears in the

²³ I take it that the problem is roughly the same on both interpretations: a circular analysis fails, because it presupposes an understanding of the concept that it sets out to explain, and similarly a metaphysical account that is uninformative fails because it does not provide the kind of explanation that it is supposed to establish: in our case, showing that reasons for attitudes are explanatorily prior to value.

²⁴ Rabinowicz / Rønnow-Rasmussen (2004), p. 414ff.

intentional content of that attitude as the feature for which its object is being admired.”²⁵ The dual role solution has nonetheless received much attention, and a number of suggestions to develop it have been advanced in the recent literature.²⁶

I am not going to discuss them in any detail here though. One reason for not doing this is that I believe that the problem has not been identified clearly enough yet (I’ll explain below). I want to approach the issue in a somewhat different way that will also allow us to place it within a broader context.

The WKR problem is a problem only for reasons for attitudes. There are no ‘wrong reasons’ for actions²⁷ – or so I will argue. Once we see why this is, the solution will take a different shape (I think).

2. *Reasons for actions and reasons for attitudes*

Only the reasons for attitudes disjunct of BPA gives rise to the WKR problem. Let me firstly explain why, and then, secondly, show how we can take our lead to solving the problem from this observation.

How do the evil demon examples work out if we focus on reasons for actions? Assume that the evil demon orders you to express your admiration for him (or for a saucer of mud) by bowing three times, or he will torture you. Clearly you now have a *pro tanto* reason to bow three times and do whatever else it takes to express admiration, and this reason is not of the wrong kind: it is a typical instrumental reason. If BPA+ were phrased in terms of reasons for action only, there would be no problem: the value of ϕ -ing (bowing, in our

²⁵ Ibid, p. 419.

²⁶ Olson (2004); Rabinowicz and Rønnow-Rasmussen (2006); Lang (2008).

²⁷ Pace Schroeder (forthcoming) who recently, to my mind unconvincingly, argued that there are.

case) may consist in ϕ -ing's having other properties (e.g. being necessary to averting the demon's wrath) that give you a reason to ϕ . The value in question is instrumental value. Thus, there is no problem with giving a buck-passing account of the value that relates to reasons for action – it is quite straightforward.²⁸

But BPA has sometimes been suggested as an account of 'final value': 'being valuable for one's own sake'. The account sets out to capture reasons to admire something for its own sake. Isn't it problematic then that the value of the action is instrumental value? But actions typically have instrumental (rather than final) value: their point is to achieve something, or bring something about. Yet there are exceptions: some actions may have final value (such as doing one's duty²⁹, acting kindly or whistling a tune). How about those actions then? If an evil demon orders you to do your duty (or else he'll torture you), you don't have a wrong reason – you just have an additional one. There may be a reason for doing your duty because of the final value of doing so, and there is now a further reason, because the demon orders you to do it, and you will be avoiding the punishment if you obey. This is no more problematic than, say, the reason to do your duty because if you don't you may lose your job. Perhaps the reason to do your duty because doing so has final value³⁰ is a sufficient reason, but there can nonetheless be additional reasons that are unrelated to the final value of so acting. As above, the evil-demon-reason is simply a typical instrumental reason. There is instrumental value in doing whatever it is that you are ordered to do, and there is no

²⁸ Parfit (2001) agrees that BPA is particularly well suited to account for instrumental value of this kind.

²⁹ At least, according some deontological accounts thereof.

³⁰ I assume that it does for the sake of argument.

particular problem with giving a buck-passing account of this kind of value – I therefore don't see a reason for restricting the account to final value.

But what if the demon ordered you to do your duty for its own sake (or because of a sense of duty)? Again, there is no problem, since you have a reason to do that anyway. And finally, what if the demon ordered you to express your admiration of *his* final value, or to express that you admire him for his own sake? You may feel a bit pressed. But it is not because the reason is of the wrong kind, but because complying with it involves a kind of pretense or lying. You don't believe that the demon has 'final value' – but you could try to play-act. If the order should be understood more stringently, as involving that you act out of the right attitude (and being a demon, your torturer will know if you don't), the problem is with having the attitude, not with acting in a certain way.

I cannot at this point offer a general explanation why reasons for actions do not give rise to the WKR problem, but I will be able to do so by the end of our journey (in II.7). Bear with me.

Let us now look more closely at the WKR problem with reasons for attitudes. If the demon orders you to admire him, or else he'll torture you, the reason is of the wrong kind, because, while you now have a reason to admire the demon, this is not because admiring him is a fitting response to his value – *ex hypothesi*, being evil, the demon has no value.³¹

This description brings back the circularity worry: we know independently that the demon is 'evil', and therefore the reason to admire him must be of the wrong kind: it is of the wrong kind because it does not

³¹ And the demon does not even instrumental value – it is only acting in a certain way that does.

relate to his value. This explanation is suspicious if offered by the buck-passer. If our only way of ascertaining that certain reasons are 'the wrong reasons' would be that our independent understanding of value allows us to characterize them as 'wrong' from the perspective of the analysis, one may begin to wonder why the analysis seemed tempting to begin with. It seems that we already understand value, independently of the analysis, and furthermore, this understanding is presupposed by the analysis.

But since this worry is couched in epistemic terms, the buck-passer can shrug it off: buck-passing is not offered as an account of our knowledge of values.³²

Or, alternatively, the proponent of BPA may reply that the lesson that the 'evil demon' recipe for generating counterexamples teaches us is a different one: the disvalue of the evil demon is stipulated, and thus does not show that we have epistemic access to value which is independent of our knowledge of reasons. The examples, rather, show that there could be reasons for admiring *anything* – because there are no limits to what the evil demon can order you to admire. But we know that it is not true that just anything is of value. Hence some of the reasons must be of the wrong kind. While the

³² Interestingly, Wallace (forthcoming) argues that the buck-passing view ought to be defended as a view about the epistemology, rather than the metaphysics, of values. He observes that we are sometimes unsure whether a certain feature or property is evaluative or normative. He suggests that if we wonder whether (e.g.) the pleasantness of a resort is a normative property, the question to pose and answer is whether it provides reasons. Thus our knowledge of reasons may be prior to our knowledge of values. However, in the current context this approach would probably fail. If, in doubt whether being the order of an evil demon is an evaluative property, you ask 'does it provide reasons?', the answer is 'yes'. Why then presume that the reason in question is a 'wrong reason', if you don't have independent knowledge of value? – I do not doubt Wallace's observation nor his suggestion that we 'test' the normativity of alleged evaluative properties by trying to figure out whether they provide reasons, by the way. All this seems entirely accurate to me. But I am doubtful that it should be seen as underwriting a buck-passing account of values.

examples make this clear by stipulating the evilness of the demon, the general point is independent of this.

Yet, there is a more important worry lurking behind the epistemic one. Take a neutral example: Some person, P, orders you to admire something, X, or else P will punish you. Neither P nor X are described in evaluative terms, and you don't know whether P is a good person, or X is an admirable thing. Is your reason for admiring X of the wrong kind? It seems to me that it is. We can say that much without knowing whether X has value or not: even if it does, its value does not consist in the reason that is provided by being ordered to admire it. But why is the reason of the wrong kind? Not because you know independently that X has no value – we assumed that you don't know this. X may be the most admirable thing. This, I believe, is a start towards showing what is wrong with the way WKR problem has been set up: The reason is of the wrong kind, but not because the thing in question has no value. Or, to put it differently, whether or not a reason for admiring is of the wrong kind does not depend on the value of what is to be admired.

To see this more clearly, let's take out the uncertainty (and the epistemic taint of the previous example). Imagine your benign and caring benefactor wants you to love a beautiful painting that she is going to bequeath you. You have one reason for loving the painting anyway, namely that it is beautiful, and you have a further reason because your benefactor wishes it. Your reason is not the avoidance of a torture threat in this case, but perhaps a reason of gratitude or a reason not to disappoint your benefactor. This reason you would have even if your benefactor, not being the best judge in matters aesthetic, were mistaken on this occasion, and the painting were but a poor, inept attempt at artwork. And furthermore, this second reason for

loving the painting is of the wrong kind in the very same way in which the reason to admire the evil demon is. In this example it is not because we stipulated that your benefactor is evil and has no value – to the contrary: we stipulated that she is good. But of course the object of the attitude is not your benefactor but the painting. But even if we assume that the painting is indeed beautiful and deserves to be admired, the reason of gratitude towards your benefactor for loving it is of the wrong kind. It just is not true that the goodness of a beautiful painting consists in its having other properties, such as being commended by a kindly person to whom you owe a debt of gratitude. If anything, its value consists in its aesthetic qualities – this we know *a priori*.

Thus perhaps the problem has not been identified clearly enough. It is not the problem that there can be reasons for pro-attitudes towards things which are devoid of value. Some reasons for admiring things are of the wrong kind, whether or not those things are of value or admirable.

3. Placing the problem

Let me therefore make a fresh start. Quite intuitively speaking, there are different kinds of reasons for having attitudes. If I am being promised £100 if I believe that it's sunny outside, I have a reason to believe it, even in the face of pouring rain, the reason being that it will earn me the money. But that seems to be a peculiar reason. The example has nothing to do with the buck-passing account of value and the WKR problem, since BPA is not concerned with reasons for belief. However the reason is peculiar in perhaps the same way in which wrong reasons are.

Joseph Raz has recently suggested that the mark of reasons of this kind is that they cannot be followed directly.³³ That is, I cannot form a belief that it is sunny directly in response to the reason that it will earn me a £100. I can comply with the reason, however. Perhaps I mistakenly happen to believe that it is sunny anyway; or I can undergo hypnosis, and end up with the relevant belief, thereby earning the money. In either case I would comply with the reason for having the belief. But I cannot form the belief directly, *for that reason*. We can respond directly to reasons for belief only if they are truth-related.

This is not meant to be a remark about psychology. The idea is not that we are psychologically unable to form beliefs in certain ways directly (even though that may by and large be true), but it is rather a claim about the normative structure of belief: if someone were to form a belief about the weather directly in response to finding that she'll get £100 if she does, then she must be conceptually confused (not understanding what 'it is sunny' means) or irrational.

Raz proposes to distinguish the two kinds of reasons as 'standard' and 'non-standard' reasons. As Raz sees it,

[s]tandard reasons are those which we can follow directly, that is have the attitude... for that reason. Non-standard reasons for... an attitude are such that one can conform to them, but not follow them directly.³⁴

Standard reasons for belief are truth-related reasons. But there are other reasons for having beliefs – perhaps again reasons of the evil demon variety,

³³ Raz (2009).

³⁴ Raz (2009), 40.

but also more mundane ones: believing that Iraq has bio-chemical weapons may help a politician to get along with his allies and colleagues. Hence, she has a reason for having the belief, but she can't directly follow that reason. As before, she can do so indirectly: by having the false belief anyway, by visiting a hypnotist, or by being self-deceived etc. And again, the 'can' and 'cannot' should not be understood as being supported by empirical findings about our psychology, but as claims about the nature of belief. If a person were to follow that reason directly it wouldn't show that it is a standard reason after all, but that she is conceptually confused, or self-deceived.

Reasons for affects and emotions are similar: Take reasons for admiration or fear. Perhaps it would be good to be afraid in a certain situation because it would make you more alert to changes in your environment, or perhaps it would be good if you admired your boss, because this would make it more likely that you will get promoted. Therefore you have reasons to be afraid or to admire, which are independent of danger and admirability. But you cannot be afraid or admire someone for those reasons (even though they are perfectly good reasons for having those attitudes).

The so-called 'wrong reasons' that haunt BPA are non-standard reasons for attitudes. They too are reasons that cannot be followed directly. But they can be conformed with. A person can take steps towards acquiring the attitude that the evil demon command requires, even if she cannot form it directly. Or she can just happen to have it (for no reasons, or bad reasons). The non-standard reason itself provides a perfectly 'good' reason for having the attitude, albeit one that cannot be followed. An example of John Skorupski's illustrates this nicely:³⁵ a demon orders someone to admire a

³⁵ Skorupski (2007).

lousy violin performance. She can conform with this reason in various ways, one being that, having bad taste in music, she believes that it is an excellent performance and therefore admires it. While her false belief that the performance is excellent is no reason to admire it, the demon's order (as ever backed by threat) is.

Raz identifies non-standard reasons with practical reasons – reasons to bring about or take steps towards having the relevant attitude. But the example shows that this is a bit rash: Complying with a non-standard reason may require taking action of some kind. But it need not. Non-standard reasons can on occasion give rise to practical reasons – but they are not themselves practical reasons. The reason to admire the bad concert is not itself a practical reason since it can be complied with without taking any action (simply by being a bad judge), but it is true nonetheless that the reason cannot be followed directly. Even the bad judge cannot respond directly with admiration to the demon's threat.

4. Explaining non-standard reasons

But why is it then that there are these two kinds of reasons for the attitude? Because attitudes, it seems – at least those attitudes we have been considering so far – have an inherent standard. As Raz sees it,

[r]easons are adaptive [i.e. 'standard reasons'] if they mark the appropriateness of an attitude in the agent independently of the value of having that attitude, its appropriateness to the way things are. (46).

Truth is the standard in the case of belief; admirability (or perhaps excellence in its various forms) in the case of admiration; danger in the case of fear. All those attitudes can be formed directly only in response to reasons that relate to the respective standards.

In the case of belief, the explanation why some reasons are standard reasons whereas others are not, is not difficult: the non-standard ones are unrelated to the truth of the belief. Admiration is similar: the standard reasons for admiring are those that speak to the admirability of that which is admired; the non-standard reasons are unrelated to admirability. They are, however, related to value: there is a non-standard reason to admire something if and because it would be good to have the attitude, or because the having of the attitude is of value.³⁶

5. Parfit's distinction of state- and object-given reasons

Derek Parfit discusses, but rejects, a somewhat similar distinction between state- and object-given reasons for attitudes (such as beliefs or desires).³⁷ Parfit does not introduce the distinction in the context of buck-passing. While he endorses a version of BPA, he does not address the WKR problem. This is not, I think, an oversight on his part, but it is because his brand of buck-passing doesn't have a WKR problem. Parfit's buck-passing account is only concerned with reasons for actions³⁸ and, as explained above, there is no WKR problem

³⁶ This relation of non-standard reasons to value seems to be Raz's reason for claiming that non-standard reasons are 'practical reasons'. As remarked above this doesn't seem quite right, as a non-standard reason can be complied with without taking any action.

³⁷ Parfit (2010), Appendix B.

³⁸ To my knowledge, while Parfit (2001 and 2010) endorses buck-passing, he does not spell out anywhere the proposition that describes his version of the account. But judging from the context, his claim is probably something like: 'x is good' consists in x's having other properties that provide *reasons for acting* in certain ways. Furthermore, he explicitly discusses

for reasons for actions. But let's investigate the suggestion since it may be of use to the buck-passer who wishes to include reasons for attitudes and has a problem with non-standard reasons.

Parfit rejects the idea that there could be both state- and object-given reasons for attitudes. A state-given reason for an attitude would be a reason to have that attitude because it would be good to have – it is 'grounded' in properties of the attitude.³⁹ An object-given reason is provided by properties of the 'object' that the attitude is an attitude towards. I.e. there is reason to believe that $2 + 2 = 4$, because it does. The object of the belief is ' $2 + 2 = 4$ ', and it is true. Parfit considers whether there could be a state-given reason to believe that $2 + 2 = 1$, if – the move is familiar now – an evil demon will torture you, unless you do. Parfit thinks we should reject this idea.⁴⁰ His worry is that if there were both state- and object-given reasons for forming beliefs about the sum, those reasons could compete or conflict. It would then be a sensible question to ask 'Do I, all things considered, have most reason to believe that $2 + 2 = 4$, or that $2 + 2 = 1$?' Parfit wishes to reject this question, because he imagines that, in some circumstances, the reason to avoid the demon's threat could be conclusive: you would then have most reason to believe that $2 + 2 = 1$. Since your reason to believe that $2 + 2 = 1$ is stronger than your reason to believe the alternative, we would have to conclude that,

buck-passing for instrumental value (which is important in the context of reasons for action, as I have explained in section II.2 above).

³⁹ Following Olson (2004: 297), we should perhaps say more precisely that it is the properties (or properties of consequences) of having of the attitude that provide the reason.

⁴⁰ Both Rabinowicz / Rønnow-Rasmussen (2004) and Olson (2004) reject the distinction for a different reason: Properties of the attitude can be recast as properties of the object of the attitude. Hence there is no real distinction here. (For a discussion of the many responses and counter-responses, and further suggestion along these lines, see Lang (2008).) I will not concern myself with this worry, because I believe that the introduction of Raz's distinction helps to put it to rest. The distinctive characteristic of standard reasons is not that the reason is provided by a property of the object of the attitude, but that the reason satisfies the standard that determines the appropriateness of the kind of attitude that it is.

in this situation, it is rational to believe that $2 + 2 = 1$. But it plainly isn't. It would be good if you did, but it would not be rational.

To avoid this problem, Parfit suggests that we should distinguish between reasons for having beliefs (and other attitudes) and they are all object-given, and reasons for *causing ourselves to have such attitudes* (which may be state-given). A reason for causing ourselves to have an attitude is not a reason for the attitude, and therefore there aren't state- and object-given reasons for attitudes. If we distinguish in this way, we would make it clear that the reasons respond to different questions. State-given reasons are practical reasons, and object-given reasons are epistemic reasons (at least in the case of belief). Hence, state-given reasons in this sense do not speak to the rationality of belief. They are relevant with regard to the rationality of taking certain actions.

The problem with accepting the possibility of state-given reasons for attitudes that Parfit points out is real enough. However, I think that we should reject Parfit's solution nonetheless.

First, the idea of an object-given reason is confusing. Take epistemic reasons: That it's overcast is an epistemic reason to believe that it will rain – hence, according to Parfit, an object-given reason. But the reason is not a property of the belief, or of the content of the belief, that it's going to rain. 'That it is overcast' is not a property of 'it's going to rain'. It is an epistemic reason because it is truth-related, not because it is a property of the object of the belief. Something similar holds for the math-example above: that $2 + 2 = 4$

is *not* itself a reason for believing that $2 + 2 = 4$.⁴¹ Any reason for believing this will not be a property of the 'object' of the belief.

Raz's distinction of standard and non-standard reasons is more helpful. The distinctive characteristic of standard reasons is not that the reason is provided by a property of the object of the attitude, but that the reason relates to the standard that determines the appropriateness of the kind of attitude it is.

The second reason for rejecting Parfit's approach is that it rests on too frail a distinction: the distinction between a 'reason to have an attitude' and a 'reason to cause oneself to have an attitude'. The reason to cause oneself to have an attitude could quite naturally be regarded as a reason for having that attitude. To keep the distinction in place we would have to regiment language, and make it terminological.

But even if we did, we would come up against a third problem: state-given reasons are reasons to cause oneself to have an attitude only if one doesn't have it anyway (due to mistaken beliefs, bad taste etc). Since state-given reasons can be complied with without 'causing oneself...', they are not properly captured by that phrase. (They are not, as such, practical reasons.) They are in the first place reasons to *have* the attitude, just as object-given reasons – the 'causing' comes in only as one way of acquiring it, because the attitude cannot be acquired by following the reason directly.⁴²

⁴¹ Alternatively, we could allow that the fact that it rains (etc.) is *a* reason for believing that it rains. I think that this is a mistake, but I don't have space to argue for this point. But even if we were to accept that the fact that *p* is a reason to believe that *p*, it would at best be *a* reason for the belief – most reasons will be different. Normally, the reason for believing that *p* will be the fact that *q*, and $q \neq p$. Therefore, normally the reason for believing that *p* will *not* be 'object-given' (on the understanding of the term that I explored above).

⁴² For a related comments see Rabinowicz / Rønnow-Rasmussen (2004), p. 411f.

Luckily, Parfit's distinction between reasons for causing and for having an attitude is not necessary for avoiding the problem: As we have seen in the context of Raz's discussion, standard and non-standard have very different explanations. Only standard reasons for beliefs are epistemic reasons, because only they relate to the epistemic standard. Hence they are the only ones that determine the epistemic rationality of believing. That there are non-standard reasons for believing (independently of epistemic reasons) does not undermine this. 'You have most reason to believe that $2 + 2 = 1$ ' is a statement that could mislead someone into thinking that having this belief is, in some circumstances, epistemically rational. But it would not be difficult to explain why this is not the case: the reasons for having the belief aren't epistemic reasons – hence, they cannot establish the epistemic rationality of the belief. Similarly in all other cases: 'there is most reason for admiring the evil demon' could mislead someone to conclude that the evil demon is admirable. But the explanation of the reason would make it clear that it is unrelated to admirability. That is simply what 'non-standard reason' means.

6. Buck-passing & the standard /non-standard reasons distinction

We have now found a way of distinguishing between different kinds of reasons for attitudes, which maps onto the buck-passer's distinction between right and wrong reasons. Does it then help the buck-passer? Is the story we told available her?

If the suggestion I have been exploring is correct then the wrong reasons are non-standard reasons in the sense that they cannot be followed directly. The explanation, why some reasons are standard reasons that can be

followed directly, and others aren't, is that some reasons relate to the inherent standard that governs the forming of the attitude, whereas others are provided by the value of having the attitude. The problem for the buck-passer is to explain the distinction in a non-circular way: To avoid explaining reasons for those attitudes that BPA is concerned with in terms of value. Does the standard / non-standard reasons distinction escape the circularity worry?

You may think that it does because the standard that governs the attitude needn't be evaluative at all: truth-relatedness is not (I think) an evaluative standard. It doesn't govern beliefs because it is good to have true beliefs. Even if, by and large, the point of forming true beliefs is that it is good for us to have them, there'll be some beliefs that are not good to have (for some person, in some circumstances), yet they are appropriate and conform to the relevant standard.

However, the standards of evaluative attitudes, such as 'admirability', are evaluative standards. But it doesn't follow that having the attitude is in any way good. There may be no point at all in my admiring some admirable things (e.g. if I will never have a chance to engage with them; they aren't accessible for me; or they are beyond the reach of my sensitivities, etc.). Yet, it would be appropriate. Forming attitudes in accordance with their inherent standard then needn't be of any value. At the very least we should regard it as an open question whether there is value in having true beliefs (let alone believing all that is true), or in being afraid when there is danger (let alone fearing everything that is dangerous). Having a standard reason for having an attitude is independent of the value of having that attitude.

There is value, however, in complying with non-standard reasons. There is a non-standard reason for having an attitude if and because it is good to have the attitude. While non-standard reasons are provided by the value of having the attitude, and thus relate to value in a very direct way, standard reasons only relate to whatever standard governs the appropriateness of the attitude, and even if this is an evaluative standard, it doesn't follow that having the appropriate attitude is *ipso facto* of value.

The buck-passer is not concerned with just any attitudes but only with *pro-attitudes*, like desiring, respecting, loving. Of the ones I have discussed above, only 'admiring' comes within her purview. Furthermore, her claim is not that having the attitude is of value, but that some object's, X's, being of value is the same property as there being a reason for having an attitude of a certain kind towards X, which is provided by some property of X, P, and P ≠ 'being good'. According to our suggestion, a standard reason for admiring X, is that X has some property P that makes X admirable.

Is this just a simplified version of BPA? It isn't. First, there is no suggestion that X's admirability is the same property as there being a standard reason for admiring it. There is no property identity claim involved. (I'll get back to this point in section III below.) Secondly, there is no reason to assume that P (the standard reason for admiring) couldn't be that X is (in a certain respect) good. Therefore, none of the distinctive features of BPA finds its way into the standard / non-standard reasons distinction.

Furthermore, the suggestion differs from BPA not only in detail, but also in spirit: It explains reasons for admiring in relation to an evaluative standard, admirability. It therefore abandons the buck-passer's ambition to

explain value in terms of reasons and instead explains reasons (for evaluative attitudes) by their relation to value. Therefore, as a response to the WKR problem the distinction of standard / non-standard reasons is circular in the way that the buckpasser needs to avoid. We said that the wrong reasons are reasons for admiring that cannot be followed directly because they are unrelated to the admirability of the thing in question. The buck-passer needs a distinction between right reasons - that are independent of admirability (or of value) - and the wrong reasons. The standard / non-standard reasons distinction doesn't help her. It explains both the standard and the non-standard reasons for pro-attitudes in terms of value: the relation to the evaluative standard and to the value of having the attitude, respectively.

If the explanation thus far has been convincing, then, as far as I can see, there is no way of spelling out the distinction that salvages the buck-passing account. Starting from the distinction between wrong and right reasons that the buck-passer brought to our attention led us to a distinction between reasons for attitudes that can be followed directly and those that can't which has applications beyond evaluative attitudes – in the case of reasons for belief, for instance. Thinking about the significance of the WKR problem – which is a circularity worry that arises for the FA-analysis of value and for BPA, insofar as it purports to explain value in terms of reasons for attitudes – leads me to the view that we should not try to solve the problem, but abandon BPA. Eschewing BPA, there is a distinction between different kinds of reasons for

attitudes – as described above – but none of them is the ‘wrong kind’, and there is no problem with them being of different kinds.⁴³

7. Reasons for actions & wrong reasons

I began this section by arguing that there are no wrong reasons for actions. Alternatively, we could (and perhaps should) reject the distinction as applied to reasons for actions. We explained standard reasons (‘right reasons’) as reasons that we can follow directly in forming an attitude. If the only way of forming an attitude in response to a reason would be by taking some action that results in having the attitude, the reason for doing so would be a non-standard reason. (But remember that I rejected Raz’s and Parfit’s claims that non-standard reasons *are* practical reasons.) Viewed from this angle it seems that practical reasons either do not fit into the standard / non-standard reasons distinction, or they are all non-standard reasons. I think we should accept the former claim because the latter is just needlessly confusing. The standard / non-standard reasons distinction is really concerned only with reasons for attitudes.

One reason why the distinction sits badly with reasons for action is that there is no equivalent to reasons that can be followed directly with regard to actions. After all, following a reason by taking actions is a typical instance of complying with a reason that cannot be followed directly.

⁴³ The WKR problem simply is the circularity worry for BPA – there is no other problem. In a recent response to Olson, Rabinowicz and Rønnow-Rasmussen (2006) seem to embrace the circularity. It is not clear to me what is left of the buck-passing view if it allows that reasons can be explained in terms of values.

What should we say then about a buck-passing account that is couched in terms of reasons for actions only? The following thesis would perhaps describe it:

(BPA_{PR} +) 'X is good' consists in X's having some other property, P, that provides a reason for action.⁴⁴

There could be a circularity worry here as well: it might turn out that we can explain reasons for action only in relation to value. But this worry is not a variant of the WKR problem. The problem is not that we need a non-circular explanation of a certain intuitive difference between kinds of reasons. There are no such different kinds of reasons when it comes to reasons for actions. The problem is simply to come up with a substantive account of reasons for actions that does not explain them by their relation to value (or as deriving from reasons for attitudes, lest the WKR problem will return). Thus, even if the WKR problem cannot be solved, this version of buck-passing could still be viable.

Two qualifications: it does encounter the old WKR problem in all those cases where a reason for action requires that the action is done with a certain attitude – if acting out of gratitude requires that you actually feel gratitude, for instance, the problem is back, for there could be an evil demon...

Secondly, this is neither a version nor a development of FA-analysis. Whatever can be said in its favor will be independent of FA-analysis.

What can be said in its favor?

⁴⁴ The 'PR' subscript stands for 'practical reason'.

III. Why buck-passing?

The buck-passing account of value brought back a theme that had retreated somewhat into the background of philosophical discussions:⁴⁵ the explanation of the relation of values and reasons. BPA explains this relation by putting forward two theses: BPA+ and BPA-.

As before, I will focus on the positive claim and on its metaphysical rather than its semantic interpretation, and I will now be only concerned with BPA_{PR}+, assuming that the WKR problem for BPA_{FA}+ cannot be solved. In its metaphysical interpretation BPA+ is a claim about property identity.

Why does the buck-passer suggest property identity of the particular kind that BPA+ describes? BPA+'s characteristic emphasis is on the claim that being of value consists in there being properties *other than being of value* which are reasons. It is not altogether clear why the buck-passer takes this view. One reason that has been offered is that being good itself is not a reason, i.e. BPA- is offered as a reason for accepting BPA+. As shown in the beginning, BPA- does not entail BPA+; and as shown elsewhere BPA- is false. I will therefore disregard this consideration.

Could an argument for BPA+ be grounded in some view about the relation of being good and the other properties that are alleged to provide reasons? It cannot, for two reasons: even if there is a necessary relation between goodness and the reason providing properties, such as pleasantness (as there would be, if the relation is one of supervenience, say), there wouldn't be property identity. A supervening property is not identical with any of the properties it supervenes upon. And the same holds if the relation is specification: if the other properties specify in the way in which something is

⁴⁵ Having been in the foreground especially in early 20th century ethics, as a question about the relation of the right and the good.

good. There would then be a necessary relation between the reason-providing properties and goodness, but it would not be identity.

More importantly, we would be looking for property identity at the wrong level. According to the buck-passer, goodness is a second-order property. It is not identical with any first order property (or properties). The claim is that goodness is the same property as the property of there being a reason to act or to have a pro-attitude of some kind (which in turn is a property the reason-providing properties – hence, a ‘higher-order’ property).

But there does not seem to be any (valid) argument in favour of this view, and furthermore, as has been pointed out by a number of philosophers, it begs the question against some deontological views of normative ethics.⁴⁶ Some deontologists claim that, while there is a reason to keep one’s promises (say), or not to kill an innocent person, it doesn’t follow that it would be good to act in these ways. It also begs the question against some accounts of reasons, like Williams’s internalism, which rests on the assumption that an account of practical reasons has to explain the difference in meaning and in truth conditions between the propositions ‘there is a reason to ϕ ’ and ‘it would be good to ϕ ’.⁴⁷ The same goes for value-based views of practical reasons, which explain reasons through values, claiming that there is a reason *because* acting in some way is good. The identity claim does not allow for the explanatory distance of the ‘because’. It seems that BPA (in its metaphysical interpretation) presupposes that the deontologist, the reasons internalist, as

⁴⁶ See for instance Dancy (2000), Crisp (2008). Olson (2007) suggests a reply to this objection, on behalf of the buck-passer.

⁴⁷ It is one of the express aims of Williams’s internalism to uncouple values and reasons. Take the following example from Williams (1995): “I shall presumably say, whatever else I say, that it be better if [a cruel husband] were nicer to [his wife]. There is one specific thing the external reasons theorist wants me to say, that the man has a reason to be nicer. [...] The question is: what is the difference between saying that the agent has a reason to act more considerately, and saying... that it would be better if they acted otherwise.” (39f) Williams assumes that there *is* a difference.

well as the reasons externalist who offers a value-based explanation of reasons are all mistaken; the semantic buck-passer would add that they are also conceptually confused. Both are strong claims – and, as far as I can see, the buck-passer provides no reason to accept them.

I take it that the virtue of BPA is that it raises interesting questions, in particular the question how to explain the close (perhaps necessary) relation of values and reasons. Yet, we are not quite sure how closely related the two are – various normative theories and theories of reasons allow for different kinds of distances. Assuming that none of their proponents is simply conceptually confused, can we settle the issue by going for metaphysical buck-passing? This seems to be closing off certain possibilities by fiat and without argument. Therefore, while the question is important, I don't see any reason to accept the buck-passer's answer.⁴⁸

References

- Brentano, F. (1934), *Vom Ursprung sittlicher Erkenntnis*, Leipzig: Felix Meiner.
- Bykvist, Krister (2009), 'No Good Fit: Why the Fitting Attitude Analysis of Value Fails' in *Mind* 118, 1-30.
- Crisp, R. (2005), 'Value, Reasons, and the Structure of Justification: How to Avoid Passing the Buck' in *Analysis* 65: 80-85.
- (2008), 'Goodness and Reasons' Accentuating the Negative', *Mind* 117(466): 257-265
- Dancy, J. (2000), 'Should We Pass the Buck?' in A. O'Hear, *Philosophy, the Good, the True and the Beautiful*, Cambridge: Cambridge University Press, 159-174.
- (2004), 'Enticing Reasons' in R.J. Wallace, P. Pettit, S. Scheffler, M. Smith (edd.), *Reason and Value*, Clarendon Press: Oxford.

⁴⁸ I am grateful to Joseph Raz and Michael Brady for very helpful comments.

- Danielsson, S. & J. Olson (2007), 'Brentano and the Buck-Passers', *Mind* 116: 511-22.
- Ewing, A.C. (1947), *The Definition of Good*, London: Routledge & Kegan Paul.
- Heuer, U. (2006), 'Explaining Reasons: Where Should the Buck Stop?' in *JESP* vol 1:3.
 - (2010), 'Wrongness and Reasons' in *Ethical Theory and Moral Practice*.
- Lang, G. (2008) 'The Right Kind of Solution to the Wrong Kind of Reason Problem', *Utilitas* 20:4, 472-89.
- Olson, J. (2004), 'Buck-Passing and the Wrong Kind of Reasons', *The Philosophical Quarterly* 54: 295-300
 - (2006) 'G.E. Moore on Goodness and Reasons', forthcoming in *Australasian Journal of Philosophy* 86: 525-34
 - (2007), 'Buck-Passing and the Consequentialism/Deontology Distinction' in *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz*. Ed. T. Rønnow-Rasmussen, B. Petersson, J. Josefsson & D. Egonsson; available at <http://www.fil.lu.se/HommageaWlodek/site/papper/OlsonJonas.pdf>
- Parfit, D. (2001), 'Rationality and Reasons' in Dan Egonsson, Jonas Josefsson, Björn Petterson & Toni Rønnow-Rasmussen (edd) *Exploring practical philosophy : from action to values*, Aldershot: Ashgate, pp. 17-39.
 - (2010), *On What Matters*, Appendix B, Oxford: Oxford University Press.
- Rabinowicz, W. & Toni Rønnow-Rasmussen (2004), 'The Strike of the Demon: On Fitting Pro-Attitudes and Value', in *Ethics* 114, 391-423.
 - (2006) "Buck-Passing and the Right Kind of Reasons", with Toni Rønnow-Rasmussen, *Philosophical Quarterly*, 114-20.
- Raz, J. (2009), 'Reasons: practical and adaptive' in D. Sobel & S. Wall (edd), *Reasons for Actions*, Cambridge: Cambridge University Press, 37-57.
- Reisner, A. (2010), 'Abandoning the Buck-Passing Analysis of Final Value' in *Ethical Theory and Moral Practice*.
- Scanlon, T. M. (1998), *What We Owe to Each Other*, Cambridge: Harvard University Press.
 - (2002) 'Reasons, Responsibility, and Reliance: Replies to Wallace, Dworkin and Deigh' in *Ethics* 112, 507-28.
 - (2008) *Moral Dimensions*. Cambridge MA: Harvard University Press.
- Schroeder, M. (2009), 'Buck-Passers' Negative Thesis' in *Philosophical Explorations* 12(3): 341-47.
 - (forthcoming) 'Value and the Right Kind of Reasons' in *Oxford Studies in Metaethics*.

- Skorupski, J. (2007), 'Buck-Passing about Goodness' in *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowicz*. Ed. T. Rønnow-Rasmussen, B. Petersson, J. Josefsson & D. Egonsson; available at <http://www.fil.lu.se/hommageawlodek/site/papper/SkorupskiJohn.pdf>
- Tappolet, Christine (2004), 'Through Thick and Thin: Good and its Determinates' in *Dialectica* 58 (2): 207-221.
- Väyrynen, P. (2006), 'Resisting the Buck-Passing Account of Value', in Russ Shafer-Landau (ed), *Oxford Studies in Metaethics*, vol 1, 295-324.
- Wallace, R. J. (forthcoming), 'Reasons, Values, and Agent-Relativity' in *Dialectica*.
- Williams, Bernard (1995), 'Internal Reasons and the obscurity of blame' in his *Making sense of humanity*, Cambridge: Cambridge University Press.