The Right Kind of Solution to the Wrong Kind of Reason Problem

GERALD LANG
University of Leeds

Recent discussion of Scanlon's 'buck-passing' account of value, which analyses the value of *X* in terms of agents' reasons for having certain pro-attitudes or contra-attitudes towards *X*, has generated the 'wrong kind of reason' problem (WKR problem): this is the problem, for the buck-passing view, of being able to acknowledge that there may be good reasons for attributing final value to *X* that have nothing to do with the final value that *X* actually possesses. I briefly review some of the existing solutions offered to the WKR problem, including those by Philip Stratton-Lake and Jonas Olson, and offer a new, better one, which accommodates all the relevant cases presented in the literature.

The so-called 'buck-passing' account of value was suggested in the work of Franz Brentano and A. C. Ewing, but has received fresh impetus by T. M. Scanlon's articulation of it in his *What We Owe to Each Other*. (The term 'buck-passing' is Scanlon's.) Scanlon defines his buck-passing account of value as follows:

> [T]he claim that [something is] 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.[1]

The buck-passing account arises out of Scanlon's immediate engagement with Moore's non-naturalist account of value, the details of which I do not have room to explore here.[2] But a brief remark on the philosophical *appeal* of the buck-passing account is not out of place, as it will help us to see what is at stake in the following discussion.

Perhaps the leading virtue of the buck-passing account is that, in Wlodek Rabinowicz's and Tonni Rønnow-Rasmussen's phrase, it 'demystifies value'.[3] Value properties are indirectly reduced to reason-providing properties, in that the property of being good is held, by the buck-passing account, to be simply the higher-order, non-reason-providing property of having other, lower-order reason-providing properties. As a result, the buck-passing account usefully provides for a

[1] T. M. Scanlon, *What We Owe to Each Other* (London, 1998), p. 96.

[2] See Philip Stratton-Lake and Brad Hooker, 'Scanlon versus Moore on Goodness', *Metaethics after Moore*, ed. T. Horgan and M. Timmons (Oxford, 2006), for a helpful discussion of these issues.

[3] See Wlodek Rabinowicz and Tonni Rønnow-Rasmussen, 'The Strike of the Demon: On Fitting Pro-Attitudes and Value', *Ethics* 114 (2004), p. 400.

© 2008 Cambridge University Press
Utilitas Vol. 20, No. 4, December 2008
doi:10.1017/S0953820808003282 Printed in the United Kingdom

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

conceptual or internal connection between the realm of the evaluative and the realm of the deontic, which means that we can pour all of our philosophical energy into identifying reasons, rather than into identifying values which those reasons supposedly, or hopefully, track.

Accordingly, the buck-passing account supposedly carries a twofold advantage: first, it obviates the need to explain the practical or action-guiding significance of value facts; and second, it obviates or at least mitigates metaphysical and epistemological difficulties about identifying value facts. (Note that Scanlon's passage states a semantic or conceptual claim about value. Even if this conceptual claim about value cannot be sustained, the buck-passing account might just work as a metaphysical account of value.)

Whether the buck-passing account really delivers these advantages, or whether these advantages can be appropriately secured in such a way, will not be my concern in this discussion. Instead, I want to pay concerted attention to a problem of an internal sort that has arisen for the buck-passing account in recent literature, which has become known as the 'wrong kind of reason' problem, or the *WKR problem*, for short.

I

The flavour of the WKR problem is conveyed in the following passages from Rabinowicz and Rønnow-Rasmussen, who have done the most to develop it:

> The reason for a pro-attitude towards an object may have to do with the value of that attitude itself and not with the value of the object.[4]

And again:

> The [WKR problem] . . . involves setting up a case in which no final value accrues to an object but there still are reasons to favour that object for its own sake, because such a pro-attitude has certain beneficial effects.[5]

Some of the cases exemplifying the WKR problem that have emerged in the recent literature, and to which I will be referring in this discussion, are summarily listed as follows (where, note, the phrases 'final value' and 'value for its own sake' are used interchangeably):[6]

[4] Rabinowicz and Rønnow-Rasmussen, 'The Strike of the Demon', p. 402.

[5] Rabinowicz and Rønnow-Rasmussen, 'The Strike of the Demon', p. 403.

[6] ED1 originally derives from Roger Crisp's review of *Value . . . and What Follows*, by Joel Kupperman, in *Philosophy* 75 (2000), p. 459; the other cases are taken from Rabinowicz and Rønnow-Rasmussen, 'The Strike of the Demon', and are further discussed by Philip Stratton-Lake, 'How to Deal with Evil Demons: Comment on Rabinowicz and Rønnow-Rasmussen', *Ethics* 115 (2005), and Jonas Olson, 'Buck-Passing and the Wrong Kind of Reasons', *The Philosophical Quarterly* 54 (2004). (Some of the labels are borrowed from Stratton-Lake.)

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

Evil Demon 1 (ED1)

If an evil demon commands us, on pain of a severe punishment, to value a saucer of mud for its own sake, we have a reason to do so.

Evil Demon 2 (ED2)

If an evil demon commands us, on pain of a severe punishment, to value him for his own sake, we have a reason to do so.

Evil Demon 3 (ED3)

If an evil demon commands us, on pain of a severe punishment, to value him for his own sake on account of the threat he has made to us, we have a reason to do so.

Paradox of Hedonism 1 (PH1)

We have a reason, by the lights of hedonism, to accord final value to things that do not, by the lights of hedonism, possess final value.

Paradox of Hedonism 2 (PH2)

We have a reason, by the lights of hedonism, not to accord final value to pleasure.

A couple of comments on these cases are in order before we proceed any further.

First, note that the soundness of PH1 and PH2 does not depend on the *truth* of hedonism as an account of individual well-being; the suggestion, rather, is that, as a purely formal rather than substantive theory of value, the buck-passing account should be able to accommodate these hedonist theses in ways that do not immediately condemn them as incoherent, even if hedonism turns out to be implausible.

Second, these cases are all concerned with final value, or the value that something has for its sake. Final value contrasts with instrumental value, which is the value that something has as a means to something else. (A refinement in the buck-passing account's distinction between final value and instrumental value will be mentioned in section V.)

II

All of these cases pose potentially severe difficulties for the simplest form of the buck-passing account, namely:

BPV1

X is good if and only if *X* has properties (other than its being good) that give us reason to have a certain positive attitude towards *X*.

This is because, in the problem cases sketched above, the object of judgment does have certain properties, at least in those circumstances, which make it appropriate for the agents to ascribe final value to it. The challenge for the buck-passer, then, is to find an amended or refined

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

form of the buck-passing account that withstands these problem cases. Many solutions to the WKR problem have been canvassed. I will not consider them all here, but only those solutions which bear either a significant relation to the solution I myself will favour, or else which have not received adequate attention in the literature.

The first solution, or first refinement, of BVP1 that we need to consider is this:

BPV2

X is good if and only if *X* has properties (other than its being good) that give us reason to have a certain positive attitude towards *X*, just as long as those properties are properties of the object, rather than properties of the attitude.

BPV2 takes extensive advantage of Derek Parfit's distinction between 'object-given properties' and 'state-given properties' (or 'attitude-given properties'), which Parfit explains as follows:

> Of our reasons to have some desire, some are provided by facts about this desire's object. These reasons we can call *object-given*. We can have such reasons to want some thing either for its own sake, or for the sake of its effects . . . Other reasons to want something are provided by facts, not about what we want, but about our *having* this desire. These reasons we can call *state-given*.[7]

In one way or another, any solution to the WKR problem is going to be prima facie drawn in the direction of Parfit's distinction. Since the 'wrong reasons' presented by the WKR problem involve reasons for having attitudes towards objects that, by assumption, do not warrant the ascription of final value, it is natural to seek ways of discarding reasons for having the attitudes that are not grounded in the properties of the object itself. Switching from BPV1 to BPV2 is thus a very natural first move.

But does BPV2 itself hit on an acceptable formulation of this ambition? According to Rabinowicz and Rønnow-Rasmussen, this solution flounders on the following problem:

> If a pro-attitude toward an object *X* would have a property *P*, then ipso facto, *X* has (or would have, if it existed) the property P^* of *being such that a pro-attitude toward [X] would have the property P* . . . [and] for any property *P* of the object of the attitude there is a corresponding property of P^* of the attitude itself: the property of *being such that its object has (or would have) property P*.[8]

Call the recipe for pairing object-given and state-given properties outlined in this passage the *pairing manoeuvre*, and the resulting

[7] Derek Parfit, 'Rationality and Reasons', in *Exploring Practical Philosophy: From Action to Values*, ed. D. Egonsson, B. Petersson, J. Josefsson, and T. Rønnow-Rasmussen (Aldershot, 2001), p. 21.

[8] Rabinowicz and Rønnow-Rasmussen, 'The Strike of the Demon', p. 406; emphases and lettering marginally altered.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

problem presented for the buck-passing account the *pairing problem*.

To see the pairing problem at work, consider ED1. The relevant reason-giving property of the pro-attitude, P, is the property of being such that having this attitude will shield me from punishment. The relevant reason-giving property of the saucer of mud, P^*, is the property of being such that my valuing it for its own sake is such that it will shield me from punishment. The difficulty with BPV2 that is presented by the pairing problem is that any reference to the properties of the object will *ipso facto* involve reference to the properties of the attitude, and vice versa. The pairing problem ensures that, for any property of the attitude, there will be a corresponding property of the object. But this means that BPV2 fails to deal adequately with the WKR problem.

III

In a recent article, Philip Stratton-Lake has complained that the pairing problem does not present insuperable obstacles to a form of the buck-passing account, such as BPV2, that trades on the distinction between object-given properties and state-given properties.[9] On Stratton-Lake's view, a solution such as BPV2 will be able to accommodate, of the cases listed above, ED1, PH1 and PH2. (He thinks, however, that ED2 and ED3 call for separate treatment: for more details, see section V, below.)

For illustrative purposes, let us consider ED1. As applied to this case, Stratton-Lake makes three claims against the pairing problem.

First, there is only one reason to value the saucer of mud, which can be described in two different ways, corresponding to an object-given property and a state-given property, respectively. There are not two reasons, one grounded in the object-given property, and the other grounded in the state-given property.

Second, the pairing problem, as Stratton-Lake interprets it, appears to presuppose that the object-given property revealed by the pairing manoeuvre must generate an extra reason for ascribing final value to the object – a reason which is additional to the reason generated by the state-given property. This seems to be because, if there really are two distinct properties in play, then the pairing problem must be committed to the view that there are two reasons in play as well.

Third, we would surely not describe that reason by invoking the object-given property *rather than* the state-given property; it would be 'incoherent',[10] in fact, to think that the reason for valuing the saucer of

[9] See Stratton-Lake, 'How to Deal with Evil Demons'.
[10] Stratton-Lake, 'How to Deal with Evil Demons', p. 793.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

mud for its own sake can be given by the object-given properties *rather than* by the state-given properties.

What are proponents of the pairing problem to make of these claims? With respect to the first claim, they are entitled to agree with Stratton-Lake: there is indeed only one reason in ED1, not two. Stratton-Lake actually admits this: 'The profligates and the parsimonists do not disagree about how many reasons *A* has to desire the saucer of mud'.[11] But if proponents of the pairing problem are entitled to agree with the first claim, it seems to me that they may also legitimately refuse to accept the second claim. They need say only that it cannot be denied that there is a *description* of the reason cast in object-given terms, invoking an object-given property, and that this description is as legitimate as a description of it which is cast in state-given terms, and which invokes a state-given property. This response to BPV2 still has critical teeth, since BPV2 clearly insists that the reason-giving properties should *be* object-given rather than state-given, not simply *reported* as object-given rather than state-given; so the sting of the pairing problem lies in its demonstration that the relevant reason-giving properties can be either object-given or state-given.

Stratton-Lake actually appears to concede this point, but then counters that: 'All [proponents of the pairing problem] have is the suggestion that there might be properties of objects that *mirror* properties of desires for those objects'.[12] However, I do not see how this claim is supposed to undermine the pairing problem, since the notion of 'mirroring' is merely a notational variant on the notion of pairing. What matters, if the pairing manoeuvre is sound, is that the object *has* properties which 'mirror', or are paired with, state-given properties, and that is enough to show that BPV2 is inadequate.

What about the third claim? In one sense, Stratton-Lake is clearly correct: it *is* incoherent to think that, in ED1, the reason for valuing the saucer of mud for its own sake can be given by the object-given properties *rather than* by the state-given properties. However, Stratton-Lake is guilty of overshooting on this score, for the converse claim is *also* incoherent: provided that the pairing manoeuvre is sound, it is also incoherent to claim that the reason for valuing the saucer of mud for its own sake can be given by the state-given properties rather than by the object-given properties. The truth, rather, is that the reason can be given by either type of property. These properties stand or fall together, in the sense that the reason for admiring the saucer of mud can be described either by invoking the properties of the attitudes, or the properties of the object.

[11] Stratton-Lake, 'How to Deal with Evil Demons', p. 792.
[12] Stratton-Lake, 'How to Deal with Evil Demons', p. 794; emphasis added.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

Stratton-Lake further remarks that 'this corresponding property [in the object] simply reaffirms the presence of the state-given reason. It does not provide a further object-given reason.'[13] But again, it is appropriate to reply that the state-given reason does not provide a *further* reason to the *object*-given reason either. If one elects to start by citing the object-given reason, then the state-given reason simply reaffirms the presence of that object-given reason. (Talk of 'reaffirmation' in this remark seems very similar to talk of 'mirroring'.)

It is of course true that the object-given description of the reason is bound to seem more laborious, artificial or gimmicky than the state-given description. Anyone's description of the reason for admiring the saucer of mud would surely start by invoking the state-given property, rather than the object-given property. Be that as it may, if one is not going to quarrel with the *reality* of the object-given property, then no real challenge has been presented to the pairing problem, and consequently no real ammunition has been offered to BPV2 for sidestepping it. And in this connection it is crucial to note that Stratton-Lake does *not* challenge the reality of the object-given property.[14] It is not clear whether Stratton-Lake thinks, in the light of his argument, merely that defenders of BPV2 are not forced to quarrel with the reality of such object-given properties, or whether, more positively, they should be committed to a metaphysically liberal notion of a property. (I suspect it is more likely to be the latter; as Jonas Olson notes, it would be tactically unwise of friends of buck-passing to insist on a very parsimonious notion of a property.[15])

IV

BPV2 cannot be saved by Stratton-Lake's strenuous interventions on its behalf. So where do we go from here? It would be unnatural, at this comparatively early stage, to surrender the hope that Parfit's distinction between object-given properties and state-given properties is relevant to an adequate solution to the WKR problem, even if BPV2 itself is inadequate.

Can an alternative deployment of the Parfitian distinction secure better results? This is Jonas Olson's hunch. Mindful of the pairing problem, Olson insists on a more thoroughgoing expulsion of state-given properties from the buck-passing analysis. That is, he wishes to eliminate *manqué* as well as explicit references to state-given

[13] Stratton-Lake, 'How to Deal with Evil Demons', p. 793.
[14] Stratton-Lake, 'How to Deal with Evil Demons', p. 793.
[15] See Olson, 'Buck-Passing and the Wrong Kind of Reasons', p. 299.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

properties in specifying the content of the reasons for our attitudes. His solution, in effect, can be stated as follows:[16]

BPV3

X is good if and only if *X* has properties (other than its being good) that give us reason to have a certain positive attitude towards *X*, just as long as the properties that give us reason for having that attitude towards *X* do not refer in any way, either explicitly or implicitly, to attitudes, or are not '*A*-referring'.

Though BPV3 seems promising, Rabinowicz and Rønnow-Rasmussen have alleged, in a reply to Olson, that BPV3 overshoots its task of excluding 'wrong reason' cases by supposedly excluding the following pair of '*right* reason' cases:[17]

Admirable

We have an additional reason for admiring an already admirable person if that person does not care about being admired.

Lovable

We have a reason to love a person who responds to love with love.

Leave aside the truth of the substantive commitments lying behind Admirable and Lovable – if this particular pair of cases is deemed to be unacceptable, other cases can surely be found in which final value can legitimately be ascribed to persons in virtue of their patterns of affective response. (In fact, it is easy to suspect, at first blush, that Admirable and Lovable may actually be in some tension with each other, since the final value of the former seems rooted in the agent's indifference to others' attitudes, while the final value of the latter seems rooted in the agent's strong non-indifference to others' attitudes – but I will let this pass. Perhaps certain distinctions between the particular pro-attitudes in question would ease this tension.[18]) Admirable and Lovable challenge BPV3 because, although these agents warrant the ascription of final value, they do so in virtue of their pro-attitudes.

One counter-response that Olson might toy with is this: the final value of the agents described in Admirable and Lovable consists, in part, in their dispositions to have certain pro-attitudes, rather than in the pro-attitudes themselves. If this is so, Olson's prohibition on *A*-referring attitudes is strictly consistent with the accommodation

[16] Olson, 'Buck-Passing and the Wrong Kind of Reasons', p. 299. Stratton-Lake, 'How to Deal with Evil Demons', p. 797n., alleges that Olson does not offer his own buck-passing account of final value, but it seems to me, in the form of BPV3, that Olson does exactly that.

[17] Rabinowicz and Rønnow-Rasmussen, 'Buck-Passing and the Right Kind of Reasons', *The Philosophical Quarterly* 56 (2006), p. 118.

[18] I will, however, have cause to return to Lovable in section VII, in order to diagnose a potential ambiguity in it.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

of these cases. However, the prohibition of *A*-referring properties in BPV3 may be too vaguely specified to make this distinction seem ultimately tenable. After all, the respective dispositions of the agents in Admirable and Lovable cannot be informatively specified without mentioning the pro-attitudes that feature in the content of those dispositions. Olson's general prohibition on *A*-referring properties does not, therefore, inspire much confidence that the dispositions can be saved, whilst allowing the attitudes themselves to be safely jettisoned. For this reason, it is reasonable to suspect that BPV3 is not a fully satisfactory replacement for BPV2. The hunt for a solution continues.

V

As we already know, Stratton-Lake offers a bifurcated solution to the WKR problem. Against ED1, PH1 and PH2, he sponsors a type of buck-passing view, such as BPV2, that relies squarely on the distinction between object-given properties and state-given properties. I have already suggested that this part of Stratton-Lake's argument fails to muster an adequate response to the pairing problem. However, the other part of Stratton-Lake's argument, which he mobilizes against ED2 and ED3 in particular, would be equally effective against ED1, PH1 and PH2, if it were sound in its own right. So this solution is worth taking seriously as a putative general solution to the WKR problem.

Stratton-Lake finesses the buck-passing view in the following way:[19]

BPV4

X is good if and only if *X* has properties (other than its being good) that give us reason to have a certain positive attitude towards *X*, just as long as those properties of *X* give us non-instrumental and non-derivative reasons for having that attitude towards *X*.

BPV4 states what Stratton-Lake calls a 'normative account', rather than a 'pro-attitude account', of the buck-passing view. The pro-attitude account of buck-passing, which Stratton-Lake attributes to Rabinowicz and Rønnow-Rasmussen, takes the buck-passing account of final value to be concerned squarely with what the reasons justify: if the reasons justify valuing *X* for its own sake, then *X* has final value, and if they justify valuing *X* for the sake of its effects, then *X* has instrumental value. The normative account, by contrast, is concerned with 'the reasons that justify a pro-attitude, rather than . . . the attitudes justified'.[20] And the reasons in question, according to BPV4, must be non-instrumental and non-derivative.

[19] Stratton-Lake, 'How to Deal with Evil Demons', p. 796.
[20] Stratton-Lake, 'How to Deal with Evil Demons', p. 796.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

To illustrate how BPV4 might be applied, consider ED2. In this case there is only an instrumental reason for ascribing final value to the evil demon: without the instrumental reason of avoiding the punishment by adopting the admiring attitude toward the demon, the agent would have no reason at all for valuing him for his own sake. BPV4's restriction of reasons to non-instrumental reasons therefore allows us to avoid having to say that the evil demon in ED2 possesses final value.

Stratton-Lake admits that ED3 is a harder case: since the agent's reason for admiring the evil demon for his own sake is due to the very threat the demon has made to the agent, ED3 ensures that there is an alignment between the content of the agent's judgment and his reason for his arriving at this judgment. This feature of ED3 makes it difficult to say that the agent's reason for admiring the demon is merely instrumental.[21] Even so, the reason is still a derivative one, since the agent would not have any reason for admiring the demon it if he did not have reasons to value the avoidance of suffering. So, on Stratton-Lake's view, BPV4 still has the resources to classify the reason for admiring the demon as a wrong reason.

However, such a prohibition on derivativeness is going to make BPV4 unable to cope with certain 'right reasons' cases. Here is one possible case:

Reformed Demon

The evil demon comes to realize the errors of his ways, sincerely repents, compensates those he has threatened, and devotes his life to good works and protecting the innocent against other, less enlightened demons.

It is plausible to think that the Reformed Demon warrants the ascription of final value. But this is so, in part, because of the good effects brought about by the mending of his ways – the compensation to those individuals he has threatened, the good works, and protection of the innocent against unreconstructed demons. So there is derivativeness here, of a sort that would be improperly condemned by BPV4.

More generally, and worryingly, it would seem that any object, *X*, which is good, is good in virtue of its underlying good-making features. Would BPV4 therefore prohibit the ascription of final value to *X*, on the grounds that *X*'s goodness is derived from its good-making features? Now I take it that Stratton-Lake would protest that he simply does not have this sort of derivativeness in mind, but the trouble is that he offers few clues as to what, precisely, he intends to designate by the

[21] I find Stratton-Lake's reasons for thinking that ED3 generates non-instrumental reasons rather elusive (see 'How to Deal with Evil Demons', p. 797), but I take this to be the most likely interpretation of his argument.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

terms 'derivative' and 'non-derivative'. It is therefore not clear how this *reductio* is to be avoided. The WKR problem remains unsolved.

VI

In more recent work, Olson, together with Sven Danielsson, has offered a distinct Brentano-influenced solution to the WKR problem.[22] This solution trades upon the distinction between 'holding-reasons' and 'content-reasons'. Danielsson and Olson define a holding-reason as a reason for having an attitude, whilst a content-reason is defined as a reason for the correctness of that attitude.[23] Content-reasons imply holding-reasons, but the implication need not hold in the other direction. It is this asymmetry of implication between holding-reasons and content-reasons which Danielsson's and Olson's solution exploits. With this background in place, Danielsson's and Olson's solution can be stated as follows:

BPV5

X is good if and only if *X* has properties (other than its being good) that give us reason to have a certain positive attitude towards *X*, just as long as those properties of *X* sustain content-reasons, not just holding-reasons, for having that attitude towards *X*.

BPV5 strikes me as unsuccessful. I shall make two comments on it. The first concerns a crucial evasiveness in the notion of 'correctness'. Danielsson and Olson admit that the notion of correctness is a 'technical philosophical notion' which is analogous to the notion of truth.[24] But they do not attempt any more detailed specification of what the correctness of an attitude consists in. As far as the WKR problem is concerned, this reticence is unacceptable. For the notion of 'a correct attitude towards *X*' can only be glossed, in this context, as 'an attitude towards *X* which is focused only on the properties of *X* which sustain right reasons, not wrong reasons'. Correctness is a place-holder, in other words, of which putative solutions to the WKR problem have to supply a detailed and precise specification. That is just what BPV5 fails to do.

Danielsson and Olson excuse themselves for not spelling out the notion of correctness any further by remarking that discussions of truth may sometimes permissibly prescind from any precise analysis – whether minimalist, deflationary, or some rival account – of the metaphysical nature of truth.[25] But this is unconvincing, for all these

[22] See Sven Danielsson and Jonas Olson, 'Brentano and the Buck-Passers', *Mind* 116 (2007). (As will be obvious, the reference they make to the present article is to an earlier, unpublished incarnation of it.)

[23] Danielsson and Olson, 'Brentano and the Buck-Passers', pp. 514–15.

[24] Danielsson and Olson, 'Brentano and the Buck-Passers', p. 516.

[25] Danielsson and Olson, 'Brentano and the Buck-Passers', p. 516.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

metaphysical accounts of truth presuppose a secure grasp of the *extension* of the predicate 'true'. For the reasons that have already been rehearsed, a similarly secure grasp of the extension of the 'correctness' of an object's good-making properties is just what is in question when we consider the WKR problem. So any evasiveness in the precise analysis of the 'correctness' of good-making properties will be failing to provide what the WKR problem requires.

Second, BPV5 strikes me – and this is something that Danielsson and Olson are particularly concerned to deny, since they deliberately distance themselves from it – as being uncomfortably close to John Skorupski's favoured solution to the WKR problem.[26] Skorupski's solution to the WKR problem rests upon a distinction between evaluative reasons and practical reasons. Evaluative reasons are defined by Skorupski as reasons for feelings and emotions, while practical reasons are defined as reasons for action. Skorupski's solution to the WKR problem – as applied to ED2, for example – is that there is only a practical reason, but not an evaluative reason, for admiring the evil demon. More precisely, the practical reason in ED2 is a reason to bring it about that one admires the evil demon.[27] But the existence of that practical reason is perfectly compatible with the non-existence of any evaluative reason to admire the evil demon.

Danielsson and Olson criticize Skorupski's solution on the grounds that, if there is a reason to bring it about that one admires the evil demon, then it seems difficult to deny that the demon has properties which make it appropriate to bring it about that one admires him.[28] In this way – leaning implicitly on the pairing manoeuvre – the practical reasons for admiring the demon imply evaluative reasons for admiring the demon. Contrary to what Skorupski suggests, there is two-way traffic between the two sets of reasons.

This criticism of Skorupski's solution seems sound. But Danielsson and Olson seem not to realize that it applies with equal force to their own solution. If there is a holding-reason for regarding the demon with admiration, then, by the pairing manoeuvre, there will be a content-reason grounded in those properties of the demon which explain why the holding-reason obtains. The pairing manoeuvre achieves two-way traffic between Danielsson's and Olson's holding-reasons and

[26] See John Skorupski, 'Buck-Passing about Goodness', *Hommage à Wlodek: Philosophical Essays Dedicated to Wlodek Rabinowicz*, ed. D. Egonsson, J. Josefsson, B. Petersson and T. Rønnow-Rasmussen (Lund, forthcoming); available at <www.fil.lu.se/hommageawlodek>.

[27] See Danielsson and Olson, 'Brentano and the Buck-Passers', p. 513.

[28] Danielsson and Olson, 'Brentano and the Buck-Passers', pp. 513–14.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

content-reasons every bit as much as it achieves two-way traffic between Skorupski's practical reasons and evaluative reasons.[29]

Now Danielsson and Olson consider the worry that BPV5 is uncomfortably close to Skorupski's solution, but they dismiss it on the grounds that BPV5 rests upon a distinction – that is, the distinction between holding-reasons and content-reasons – *within* the space of reasons for attitudes, whereas Skorupski's solution invests in a distinction between reasons for attitudes and reasons for actions.[30] But this difference is simply too slim to permit BPV5 to maintain a respectable distance from Skorupski's solution. Their criticism of Skorupski's solution, after all, might be put in the following terms: the distinction between reasons for attitudes and reasons for actions does not repay Skorupski's investment in it, because, due to the pairing manoeuvre, practical reasons are only nominally different from evaluative reasons. The pairing manoeuvre sees to it that, for each practical reason to favour the demon in ED2, there is a corresponding evaluative reason to favour the demon in ED2. Practical reasons in Skorupski's solution are thus easily transmuted into evaluative reasons – evaluative reasons of the wrong kind, of course. But if this is why Skorupski's solution is deemed to be unsuccessful, there can be no good reason why BPV5 is entitled to refuse association with Skorupski's solution by insisting that the distinction between holding-reasons and content-reasons is a distinction within the space of reasons for attitudes. Even if there are taxonomical differences between BPV5 and Skorupski's solution as they are initially described, it does not follow that those solutions may prove to be unsuccessful for very similar reasons.

VII

The final solution to the WKR problem that I will consider is this:

BPV6

X is good if and only if *X* has properties (other than its being good) that give us reason to have a certain positive attitude towards *X*, just as long as those properties of *X* that give us reason to have that attitude towards *X* would still be reason-giving in the absence of the benefits to us of having that attitude towards *X*.

BPV5 avoids the problems of Stratton-Lake's BPV4, which relies on a problematic notion of 'non-derivativeness'. And it is not troubled by

[29] Thus I think Skorupski is perfectly correct to suppose that Danielsson's and Olson's criticism of his solution applies to their own: see 'Brentano and the Buck-Passers', p. 518n.

[30] Danielsson and Olson, 'Brentano and the Buck-Passers', pp. 517–18.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

the pairing problem, for BPV6 delivers perfectly satisfactory solutions to the problem cases considered thus far.

BPV6 is plainly untroubled by the trio of ED1, ED2 and ED3: in each of these cases, there would be no reason for ascribing final value to either the saucer of mud or the evil demon were it not for the punishment-avoiding advantages of holding those pro-attitudes. What about PH1 and PH2? Again, there is no problem. According to hedonism, there would be no reason for ascribing final value to non-mental entities, were it not for the pleasure-producing effects of doing so. BPV6 therefore allows the hedonist to say that an agent may have reasons to accord final value to an object, even though that object, by the lights of hedonism itself, may not actually possess final value. BPV6 does not deny, of course, that the personal benefits caused by one's *experience* of *X* may be reason-giving, as it might be if one's experience of *X* produced pleasure. That is a central commitment of hedonism (though one does not have to be a hedonist to agree with hedonists, in at least a large range of cases, that the pleasure caused by experiencing *X* will give us a reason for experiencing *X*). What BPV6 specifically discounts are the personal benefits caused by, or associated with, the *attitude* towards *X* that constitutes acknowledgement of *X*'s final value. (I shall return to this point shortly.)

What about right reasons cases? BPV6 has no difficulty with Reformed Demon, since it can accommodate the fact that the sort of derivativeness operating in this case is perfectly compatible with the Reformed Demon's possession of final value. Neither does BPV6 have any problem with Admirable and Lovable. Though the final value of the agents in Admirable and Lovable has plenty to do with *their* pro-attitudes, or with their dispositions to have those pro-attitudes, it has nothing to do with *my* pro-attitudes, or with the benefits conferred upon me by having my pro-attitudes.

However, we would do well to pause at this point to consider Lovable in a little more detail. The problem with Lovable is that it is potentially ambiguous. As stated earlier, it claims that we have reason to love an agent who responds to love with love. The danger for BPV6, then, is that the final value of a lovable agent, whom we might call *Alison*, calls for a certain response from us – that of loving Alison – which will in turn elicit a response from her that, plausibly, benefits us. (It is plausible, or at least not incoherent, to suppose that being loved by someone we love benefits us.) But if this is so, then Alison's final value cannot be assessed by discounting the benefits conferred upon me by loving her.

This objection is confused, for there are two kinds of loving attitude we may hold towards Alison. One kind of loving attitude is concerned with the attitude it is appropriate to hold towards her in virtue of,

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

specifically, her final value – this is, in effect, a disinterested sense of loving. Another sort of reason is concerned with the reasons there are for *falling in love* with her – on this score, her high responsiveness to love may favour my romantic overtures towards her. (For example, I will not be running the risk of loving her unrequitedly.)

It should be clear that BPV6 is concerned with the first, disinterested, non-romantic sense of 'loving'. If I love or admire Alison in *this* sense, am I thereby benefited? That is not at all obvious. It is far from obvious whether she will return this sort of admiration with love; and, even if she does, it is far from obvious that I will be thereby benefited.

With respect to the second kind of loving, it is much more plausible to say that I am benefited. (I'm in love! And she loves me!) But that is not the sort of attitude in which BPV6 is interested, as I have explained.

VIII

A potentially new problem arises at this point.[31] Consider this case:

Evil Demon 4 (ED4)

If a necessarily existing, necessarily evil demon commands me, on pain of a severe punishment, to value him for his own sake, I have a reason to do so.

In ED4, the evil demon is construed as existing, and existing *as* an evil demon, in every possible world, and as presenting me with this threat in every possible world in which I exist. (Of course, ED4 does not assume that *I* exist in every possible world.)[32] The difficulty presented by ED4 for BPV6 is that it appears to rob BPV6 of the modal elbow-room required to draw the distinction between right reasons and wrong reasons. By assumption, in every world in which I exist, the evil demon is presenting me with the threat of punishment, and so there are no possible worlds to which the counterfactual test of BPV6 can be applied.

So does BPV6 falsely imply that the reasons I have, in ED4, for admiring the evil demon are right reasons, rather than wrong reasons? By no means; for the absence of possible worlds to which the counterfactual test of BPV6 can be applied implies simply that we are not in a position to say that this evil demon is good. If BPV6's test for goodness in ED4 cannot be applied, then BPV6 is not in a position to attribute goodness to the evil demon. Neither is BPV6 in any sense committed to the residual presumption, in the event of our inability to run the counterfactual test, that the evil demon is to be considered good in virtue of the reasons there are to admire the evil demon.

[31] Thanks to Andy McGonigal for pressing this possibility on me.

[32] In presenting ED4, I have switched from the first-person plural to the first-person singular, but this does not leave us with any troubling inconsistency with the other cases I have considered.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

IX

In conclusion, BPV6 certainly has something in common with the other attempted solutions, since, like all these other solutions, it is in the business of trying to discard the irrelevant state-given properties in order to privilege the relevant object-given properties. Unlike those other solutions, however, BPV6 gives us a sufficiently precise recipe to deal with all these problem cases whilst not departing from the permitted buck-passing ontology, an ontology that involves only properties of objects, properties of attitudes, and reasons for having certain attitudes to objects that are grounded in the properties of either the objects or the attitudes or both. I submit that, on the evidence so far, BPV6 is the right solution to the WKR problem.

X

For all that, is BPV6 the right *kind* of solution to the WKR problem? Might it be guilty of meeting the letter but not the spirit of the problem? To raise some doubts about the status of, not just BPV6, but also all the other solutions that have been canvassed here, I close with a further worry, and the sketch of a response – or a couple of responses – to that worry.

The worry can be characterized by distinguishing between two problems that the WKR problem might be thought to present for the buck-passing account.

One way of approaching the WKR problem is to treat it simply as what we might call a *sorting problem*: this is the problem of finding a way of couching the 'wrong reasons'/'right reasons' distinction in buck-passing terms, or using only those items that belong to the buck-passing ontology, as this ontology was described above. As I have already argued, I believe that BPV6 musters an adequate response to the sorting problem for the buck-passing account.

However, the WKR problem seems to me to be getting at some other problem as well, which I will call the *directionality problem*. The directionality problem is the problem of explaining why it is necessary, or at least so much easier or natural, to start with the left-hand side of the buck-passing biconditional than it is to start with the right-hand side. As this discussion, with its several twists and turns, may attest, it takes considerable effort to reach an articulation of the buck-passing account that matches the right-hand side with the left-hand side. What this may suggest is that we have independent intuitions about goodness, and about the location of goodness, which involve no consultation with the items that figure in the buck-passing ontology.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

The problem has been to articulate those independent intuitions about goodness in a way that respects the buck-passing ontology. But this means that it is our intuitions about value that are driving the formulation of the right reasons/wrong reasons distinction, rather than the other way round.

The worry, then, is that we are simply *not* buck-passers about value. We do not think about value in buck-passing terms. In solving the WKR problem, we have been using the category of value to illuminate the category of reasons. It follows that at least some of the advantages claimed for the buck-passing account, such as that of achieving epistemic traction on values, will now ring hollow.

Two replies to the directionality problem seem available. According to the first of them, we should distinguish between metaphysical worries and epistemological worries. In metaphysical terms, if values can be reduced in the recommended buck-passing way to reasons, then there is just one category of normative items to worry about, rather than two. How we secure epistemic access to reasons raises problems that the buck-passing account does not pretend directly to address. As far as it goes, I think that this is a fairly satisfactory reply, particularly when it is remembered that, as a purely formal model of value, the buck-passing account is not in a position to address questions arising over our identification of substantive theories of value. Presumably the identification of substantive theories of value, and so the selection of substantive reasons that constitute those theories, will be facilitated by the sort of theorizing that belongs to normative ethics.

This first reply will also allow the other advantage claimed for the buck-passing account, which is that of providing for an internal connection between value and action-guidingness, to remain intact. Whether buck-passers are justified in ensuring the existence of such a connection by defining value in their preferred way is, however, a moot point.

A second, bolder response to the directionality problem, which I actually prefer, is to deny that independent thoughts about value, on the left-hand side, really are driving the refinements that are required on the right-hand side. Indeed, it might be said that the pairing problem, with the proliferation of evaluative properties that it sets in motion, necessitates a consultation of our independent thoughts about *reasons*. After all, the evolution of the right solution to the WKR problem, from BPV1 to BPV6, has been primarily facilitated by our reflection on which reason-giving properties are relevant – it is this reflection which has enabled us to find ways of matching those reason-giving properties to the left-hand side of the biconditional.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press

In conclusion, though the buck-passing account of value may face other difficulties, it is not troubled by the WKR problem in either of its guises: it is defeated neither by the sorting problem nor by the directionality problem.[33]

g.r.lang@leeds.ac.uk

[33] Many thanks to Kent Hurtig, Arto Laitinen, and the participants in a graduate colloquium on practical reason held in the Philosophy Department at the University of Jävyskylä in December 2005, for very helpful and constructive comments on an earlier version of this material. Thanks also to Ulrike Heuer, Andy McGonigal, Andrew Reisner and, especially, Jonas Olson, for useful exchanges about these issues and/or comments on later versions of this article.

https://doi.org/10.1017/S0953820808003282 Published online by Cambridge University Press