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

JONAS OLSON
Stockholm University

The so-called *Wrong Kind of Reason* (WKR) problem for Scanlon's 'buck-passing' account of value has been much discussed recently. In a recent issue of *Utilitas* Gerald Lang provides a highly useful critique of extant proposed solutions to the WKR problem and suggests a novel solution of his own. In this note I offer a critique of Lang's solution and respond to some criticisms Lang directs at a Brentano-style approach suggested by Sven Danielsson and me.

I. INTRODUCTION

According to T. M. Scanlon's 'buck-passing' account of value, to be valuable is not to possess value as a simple and unanalysable property, but rather to have other properties that provide reasons to take up an attitude in favour of its owner or against it, or to act in certain ways in regard to it.[1] A serious challenge to the buck-passing account is that there appear to be scenarios in which there is reason to take up an attitude in favour of an object or against it, even though this has no bearing on the value of that object. This challenge has become known as the *Wrong Kind of Reason* (WKR) problem.[2] Gerald Lang has recently provided a highly useful critique of extant solution proposals to the WKR problem and has also suggested a novel solution of his own.[3] In this note I shall offer a critique of Lang's solution (section II) and respond to some criticisms Lang directs at the Brentano-style approach recently suggested by Sven Danielsson and me (section III).[4]

II. LANG'S SOLUTION: EXPOSITION AND CRITIQUE

The WKR problem has been extensively discussed in recent literature and as Lang notes the problem can be illustrated in a variety of ways.[5]

[1] T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass., 1998), pp. 95–100.
[2] The term was coined by Wlodek Rabinowicz and Toni Rønnow-Rasmussen, 'The Strike of the Demon: On Fitting Pro-Attitudes and Value', *Ethics* 114 (2004), pp. 391–423.
[3] Gerald Lang, 'The Right Kind of Solution to the Wrong Kind of Reason Problem', *Utilitas* 20.4 (2008), pp. 472–89.
[4] See Sven Danielsson and Jonas Olson, 'Brentano and the Buck-Passers', *Mind* 116 (2007), pp. 511–22.
[5] Lang, pp. 473–4.

© 2009 Cambridge University Press
doi:10.1017/S0953820809003501
Utilitas Vol. 21, No. 2, June 2009
Printed in the United Kingdom

https://doi.org/10.1017/S0953820809003501 Published online by Cambridge University Press

Here it suffices to give just one example. Suppose that a malicious and powerful demon threatens to inflict severe pain on us unless we favour him. That seems to provide strong reason for us to favour the demon, but it is obviously not the case that such a demon is thereby positively valuable. In other words, the reason to favour the demon in this scenario is of the wrong kind from the point of view of value analysis. But as recent debate has shown, it is notoriously difficult to formulate a satisfactory definition of the right kind of reasons. After rejecting a number of recent attempts Lang offers his own proposal, which he calls BPV6:

BPV6

X is good iff *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 the attitude towards *X*. (Lang, p. 484)

BPV6 is intuitively plausible as it seems to take care of a good deal of the examples of the WKR problem that have been suggested.[6] But BPV6 faces a number of challenges. First, it seems that BPV6 needs to be interpreted liberally if it is to take care of all examples that have been suggested. For instance, in order for BPV6 to take care of the example mentioned above we must assume that something that saves us from pain that we would otherwise have suffered benefits us. But let us grant that such a liberal interpretation of what it is for something to benefit us is tolerable in the present context. Second, it seems that BPV6 needs to be reformulated slightly in order to take care of other examples that easily spring to mind. Suppose for instance that an evil demon threatens to inflict severe pain on people on the other side of Earth, who are all strangers to us, unless we favour him. That seems to provide reason for us to favour the demon even though we would not benefit from favouring him. According to BPV6 this means that the demon is good. This is obviously unacceptable so BPV6 would have to be reformulated along the following lines:

BPV7

X is good iff *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

[6] In discussions a number of people have pointed out that an initial worry about BPV6 is that it is not clear that the notion of *benefits* is non-evaluative. If it is evaluative, the analysis becomes circular and BPV6 is a non-starter. It is not clear to me exactly what Lang means by 'benefits' but I shall grant him the charitable assumption that the notion can be explicated non-evaluatively. But as we shall see, it might well be far from easy for Lang to avoid including evaluative notions in the analysans, thereby making the analysis circular. See the discussion of BPV9 and BPV10 below.

https://doi.org/10.1017/S0953820809003501 Published online by Cambridge University Press

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, or anyone else, of having the attitude towards *X*.

I take it that such a slight reformulation is still in the spirit of the basic idea behind BPV6.

But third, we can formulate counterexamples to BPV7 that make that basic idea look much more questionable. Suppose that an evil demon threatens to create a number of individuals who suffer greatly unless we favour him. If we favour him he will not bring these suffering individuals into existence. That seems to provide reason for us to favour the demon even though no one would benefit from our favouring him. This is because one cannot benefit individuals who do not and will not exist. According to BPV7 this means that the demon is good. This is obviously unacceptable so BPV7 would have to be reformulated along something like the following lines:

BPV8

X is good iff *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 even if our positive attitude towards *X* would benefit no one, and even if it would not prevent occurrences of suffering.

BPV8 is more complex than its precursors, but perhaps not objectionably so. But further problems lie ahead.

Fourth, suppose that an evil demon tells us that unless we favour him he will erase completely something that many people cherish deeply – such as the works of Shakespeare – from collective memory. That seems to provide reason to favour him. But since the demon threatens to erase completely all works of Shakespeare from collective memory our favourable attitude towards him will benefit no one and it will not prevent suffering from occurring. According to BPV8 this means that the demon is good. This is obviously unacceptable so BPV8 would have to be reformulated. How? Well, Lang could suggest that eradicating all works of Shakespeare from collective memory would make our lives less rich and on this basis he could offer the following reformulation:

BPV9

X is good iff *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 even if our positive attitude towards *X* would benefit no one, and even if it would not prevent occurrences of suffering, and even if it would not prevent our lives from becoming less rich.

https://doi.org/10.1017/S0953820809003501 Published online by Cambridge University Press

BPV9 is more complex than its precursors but a more striking problem is that the claim that something would make our lives more or less rich is clearly an evaluative claim. Since BPV9 is an analysis of value that involves an evaluative notion in the analysans it is circular.[7] But perhaps there is another option. The reformulation of BPV9 could mention the supervenience base of the evaluative property appealed to in the analysans, rather than the evaluative property itself. This would give us BPV10:

BPV10

X is good iff *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 even if our positive attitude towards *X* would benefit no one, and even if it would not prevent occurrences of suffering, and even if it would not prevent all works of Shakespeare from being eradicated from collective memory.

But it is obvious that only lack of imagination stops us from conjuring up demons that threaten to eradicate from collective memory the works of Plato, Aristotle, Dostoevsky, and what have you. In order to account for this possibly infinite series of counterexamples the reformulation of BPV10 would have to take the form of a hopelessly complex conjunction.

III. A BRENTANO-STYLE APPROACH: OUTLINE AND DEFENCE

The gist of the Brentano-style approach suggested by Sven Danielsson and me is that the buck-passer should distinguish between two types of reasons: *reasons for the correctness of an attitude* (what we call *content-reasons*) and *reasons to have an attitude* (what we call *holding-reasons*). The former type of reasons trades on Brentano's notion of correctness which applies to conative attitudes. Correctness is a normative notion but crucially, from the fact that some attitude would be correct it does not *follow* that we ought to have that attitude (i.e. that there are holding-reasons for that attitude). It is however a plausible normative assumption that we normally ought to have correct attitudes. Hence we suggested that content-reasons for an attitude normatively imply

[7] Krister Bykvist suggested that the evaluative notion might be that of *value-for* rather than *final value*; what makes our lives richer has value *for* us but need not be valuable *for its own sake*. Such a manoeuvre would analyse final value partly in terms of value-for so BPV9 would not be an analysis that reduces values to reasons. One could of course attempt an analysis of value-for in terms of reasons but that would only postpone the problem for variants of the WKR challenge would arise for such analyses too. It is obvious that on pain of circularity such analyses cannot appeal to value-for or final value in order to circumvent WKR problems.

https://doi.org/10.1017/S0953820809003501 Published online by Cambridge University Press

holding-reasons for that attitude.[8] Furthermore, we suggested that holding-reasons are analysable in terms of content-reasons: to say that there is holding-reason for some attitude, a, is to say that there is content-reason to favour the occurrence of a, or possibly that there is content-reason to disfavour the non-occurrence of a.[9]

Given the Brentano-style approach, the buck-passer can maintain that an object, X, is valuable iff there is content-reason to favour X. This means that in the example given in the previous section there is holding-reason but no content-reason to favour the demon. Favouring a demon that threatens to inflict severe pain on us if we do not favour him would not be correct, but if the threat is credible we ought nevertheless to do so. That there is holding-reason to favour the demon means that there is content-reason to favour favouring the demon, which is the right result since favouring the demon is instrumentally valuable in that it shields us from pain.

Lang criticizes the Brentano-style approach on two counts. First, he argues that the notion of correctness is too evasive and, second, that the approach is only 'nominally different' from a suggestion by John Skorupski that Danielsson and I aim to reject.[10] I begin with the second criticism.

Skorupski's basic move is to distinguish not between types of reasons for attitudes but between reasons for actions (what he calls *practical reasons*) and reasons for attitudes (what he calls *evaluative reasons*).[11] Utilizing this distinction, the buck-passer can say that X is valuable iff there is evaluative reason to favour X.[12] The next move is to say that there can be reason to *bring it about* that one has some attitude without there being reason to *have* that attitude. Reasons to bring it about that one has some attitude are reasons for actions, or series of actions. Reasons to bring it about that one has some attitude are thus practical reasons while reasons to have some attitude are evaluative reasons. In the demon scenario in the previous section, there is (practical) reason for us to bring it about that we favour the demon but no (evaluative)

[8] Danielsson and Olson, p. 515. Lang fails to note that the implication here is normative and not conceptual. This might partially explain his misunderstanding of the Brentano-style approach to be discussed below.

[9] Danielsson and Olson, p. 518.

[10] Lang, pp. 482–4.

[11] See John Skorupski, 'Buck-Passing about Goodness', in T. Rønnow-Rasmussen, B. Peterson, J. Josefsson and D. Egonsson (eds.) [online resource]: *Hommage á Wlodek: Philosophical Papers Dedicated to Wlodek Rabinowicz* (Lund, 2007). Available at www.fil.lu.se/hommageawlodek.

[12] This may sound circular, but it isn't as long as 'there is evaluative reason to favour X' is taken to mean simply that there is reason to have a favourable attitude towards X and as long as the favourable attitude is specified without appeal to evaluative concepts.

https://doi.org/10.1017/S0953820809003501 Published online by Cambridge University Press

reason for us to favour the demon. Hence, the buck-passer need not conclude that the malicious demon is good.

Danielsson and I reject this approach. Lang takes our ground for rejection to be 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 that make it appropriate to bring it about that one admires him'.[13] But this is not in fact our ground for rejecting Skorupski's approach. (I don't think that Skorupski would be concerned to deny that the demon has properties which make it appropriate to bring it about that one admires him.) Our complaint was that it is difficult to find an independent rationale for making the required distinction between reasons to bring it about that one has an attitude and reasons to have an attitude. In the demon scenario, there is a strong intuition to the effect that there is reason to favour the demon (that is part of the explanation why the WKR problem is so troubling); after all, the reward (of being spared from punishment) is tied to actually favouring the demon and not merely to bringing it about that one favours him.[14] So in so far as there is reason to bring it about that we favour the demon this would seem to be so *because* there is reason for us to favour the demon. It is odd to say that we have reason to bring it about that we be in a state (e.g. of favouring the demon) that we have no reason to be in, and possibly even that we have reason not to be in (on the assumption that there is evaluative reason not to favour, or to disfavour, the demon and not merely no evaluative reason to favour him).

Lang acknowledges that he accepts our criticism of Skorupski (although he misconstrues it). He goes on to say, however, that the Brentano-style approach collapses into something equivalent to Skorupski's approach. This is allegedly because

> If there is a holding-reason for regarding the demon with admiration, then . . . there will be a content-reason [for regarding the demon with admiration] grounded in those properties of the demon which explain why the holding-reason obtains. (p. 483)

But here Lang simply fails to note that we state clearly that '[the fact] that there is a holding-reason for some attitude does not imply that there is a content-reason for that attitude'.[15] The fact that there is a holding-reason for some attitude, a, entails merely that there is a content-reason to favour the occurrence of a, or possibly to disfavour the non-occurrence of a, since, as has already been noted, this is what it means to say that there is a holding-reason for a. One merit of

[13] Lang, p. 483.
[14] Rabinowicz and Rønnow-Rasmussen press similar criticism against views similar to Skorupski's.
[15] Danielsson and Olson, p. 515.

https://doi.org/10.1017/S0953820809003501 Published online by Cambridge University Press

this view is that it accounts for the intuition that there *is* reason to favour the evil demon who threatens to inflict severe pain on us if we do not favour him. But since this is a holding-reason it does not imply that he is good. We are not committed to Skorupski's dubious claim that there is reason to bring it about that we favour the demon although we have no reason to favour him (and presumably also reason to disfavour him!). Again, while Skorupski's approach distinguishes between reasons for attitudes and reasons for action, the Brentano-style approach distinguishes between types of reasons for attitudes.[16]

I now turn to Lang's criticism that the notion of correctness is too evasive and that Danielsson and I fail to say anything illuminating about it. As mentioned, the notion of correctness is adapted from Brentano and when we introduced it in our article we said that according to Brentano correctness is analogous to the notion of truth. We noted that according to Brentano, correctness and truth are fundamentally epistemic notions that must be understood with reference to the notion of experiences of self-evident knowledge. We then went on to say that much as we may find it both acceptable and useful to talk about true beliefs without being precise about the nature of the truth predicate, we may also find it acceptable and useful to talk about correct conative attitudes without being precise about the nature of the notion of correctness. We remained neutral between rival metaphysical and epistemological accounts of correctness.[17]

Lang finds this unacceptable since being non-committal regarding the nature of the truth predicate 'presuppose[s] a secure grasp of the *extension* of the predicate "true"' but – he seems to suggest – we don't have such a grasp of the extension of 'correct'.[18] This last claim is contestable. I am inclined to believe that many people have a fairly secure grasp of the extension of 'correct'. For instance, I am sure that many people would agree that it would not be correct to favour demons who threaten to inflict severe pain on us if we do not favour them.[19]

I suspect, however, that Lang's real worry here is one that he develops in more detail towards the end of his article. This is what he calls the

[16] Lang recognizes this difference but claims that '[it] is simply too slim to permit [the Brentano-style approach] to maintain a respectable distance from Skorupski's' (p. 484). Given what has just been said above I do not see that the difference is 'too slim'.

[17] See Danielsson and Olson, p. 516. See also Franz Brentano, *The Foundation and Construction of Ethics* (London, 1973), pp. 127–32. *Cf.* Brentano, *The Origin of Our Knowledge of Right and Wrong* (London, 1969), §§23, 27.

[18] Lang, p. 483.

[19] Brentano is obviously a philosopher who claimed at least implicitly to have a secure intuitive grasp of correctness. C. D. Broad and A. C. Ewing made claims about the *fittingness* of conative attitudes that resemble Brentano's claims about correctness. See e.g. C. D. Broad, 'Some of the Main Problems of Ethics', *Philosophy* 21 (1946), pp. 99–117; A. C. Ewing, *The Definition of Good* (London, 1947).

https://doi.org/10.1017/S0953820809003501 Published online by Cambridge University Press

directionality problem. To put it briefly: how do we know that 'it is [not in fact] our intuitions about value that are driving the formulation of the right reasons/wrong reasons distinction'?[20] If it is, it seems that we have not succeeded in reducing values to reasons. This is a fair worry. Applied to the Brentano-style approach the worry becomes the following twofold one: first, can we really distinguish between holding-reasons and correctness-reasons for an attitude; second, can we do this in a way that does not appeal to more fundamental intuitions about value?[21]

In response to the first question, it should be acknowledged that in the demon scenarios most of us have intuitions to the effect that there is reason to favour the demon but there is also reason not to favour him. The way to account for and regiment these intuitions is to say that in such cases there are holding-reasons to favour the demon (we ought to favour him) but there is no content-reason to favour him (it would not be correct to favour him), where holding-reasons to favour the demon do not imply content-reasons to favour the demon. In response to the second question, let me point out that Danielsson and I close our article by posing this very question. We concede that 'what provides the content-reasons to favour x will be roughly coincidental with . . . the good-making properties of x'.[22] One might then ask why an analysis of value in terms of content-reasons for attitudes is at all preferable. One attractive feature of such an account that we stress is that it establishes the following plausible link between values and attitudes: necessarily, to claim that an object is valuable is to claim that a favourable attitude towards that object is (would be) correct.[23]

IV. CONCLUSION

On the basis of the critique mounted in section II above I conclude that Lang has offered the wrong kind of solution to the WKR problem. I maintain that the Brentano-style approach outlined and defended briefly in section III above remains the only extant proposal that promises to offer the right kind of solution to the WKR problem.[24]

jonas.olson@philosophy.su.se

[20] Lang, p. 488.

[21] I suspect that Lang is sceptical on this score and that this is what prompts him to say that whenever there is holding-reason to favour the demon there will also be content-reason to favour him. See above and Lang, p. 483.

[22] Danielsson and Olson, p. 520.

[23] Lang points out that it is a virtue of BPV6 and of the buck-passing account more generally that it establishes a link between values and attitudes (p. 25).

[24] I thank Krister Bykvist, Sven Danielsson, Wlodek Rabinowicz, Toni Rønnow-Rasmussen, and especially Gerald Lang for helpful comments and discussions.

https://doi.org/10.1017/S0953820809003501 Published online by Cambridge University Press