



# LUND UNIVERSITY

## Buck-passing and the right kind of reasons

Rabinowicz, Wlodek; Rønnow-Rasmussen, Toni

*Published in:*  
Philosophical Quarterly

*DOI:*  
[10.1111/j.1467-9213.2005.00432.x](https://doi.org/10.1111/j.1467-9213.2005.00432.x)

2006

[Link to publication](#)

*Citation for published version (APA):*

Rabinowicz, W., & Rønnow-Rasmussen, T. (2006). Buck-passing and the right kind of reasons. *Philosophical Quarterly*, 56(222), 114-120. <https://doi.org/10.1111/j.1467-9213.2005.00432.x>

*Total number of authors:*  
2

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# Buck-Passing and the Right Kind of Reasons

Wlodek Rabinowicz and Toni Rønnow-Rasmussen

The so-called buck-passing account of value faces what might be called the ‘Wrong Kind of Reasons’-problem.<sup>1</sup> Jonas Olson has recently suggested how to resolve this difficulty.<sup>2</sup> In this note, we argue that, despite its merits, Olson’s solution is unsatisfactory. We go on to suggest that the buck-passing account might be acceptable even if the problem in question turned out to be unsolvable.<sup>3</sup>

## *Buck-passing and the wrong kind of reasons*

Buck-passers analyse value in terms of reasons for pro-attitudes. They take the value of an object to consist in the existence of normative reasons for favouring it, with “favouring” being a place-holder for various pro-attitudes and pro-responses. In the pluralist version of this view, which we find especially attractive, different kinds of value (admirability, desirability, etc) correspond to different kinds of pro-attitudes that might be called for by the properties of the object. Depending on whether the properties in question provide reasons to favour the object for its own sake or for the sake of its consequences or applications, the value of the object is taken to be final or instrumental, respectively. The characteristic feature of this proposal is that the reasons for favouring are to be found in what makes the object valuable (i.e. in its value-making properties) and not in its value itself: In this sense, the ‘buck’ is being passed from the latter to the former.<sup>4</sup>

The obvious problem with this analysis is that examples can be found in which there are reasons to favour an object that lacks value. This will be the case when it is the favouring, and

---

<sup>1</sup> See our paper “The Strike of the Demon - On Fitting Pro-attitudes and Value”, *Ethics* 114 (2004), pp. 391-424. Cf. also Justin D’Arms and Daniel Jacobson, “Sentiment and Value,” *Ethics* 110, (2000): 722-48.

<sup>2</sup> “Buck-Passing and the Wrong Kind of Reasons”, *The Philosophical Quarterly* 54 (2004), pp. 295-300.

<sup>3</sup> We are grateful to John Broome, Krister Bykvist, Christian Piller, Michael Zimmerman and not least to Jonas Olson himself, for very useful comments and discussion. We are also indebted to anonymous referees. The work on this paper was supported by a research grant from The Bank of Sweden Tercentenary Foundation.

<sup>4</sup> Cf. T. M. Scanlon, *What we Owe to Each Other*, Harvard University Press, Cambridge, Mass., 1998, p. 97.

not the object itself, that has value, final or instrumental one, or when such favouring is required on deontological grounds. In our recent paper on the subject, we refer to this difficulty as the ‘Wrong Kind of Reasons’-problem, or the *WKR-problem*, for short.

Several years ago, Roger Crisp gave an example of this kind.<sup>5</sup> A malicious demon will impose a severe punishment unless we desire a saucer of mud. Or, to vary this example, the demon will punish us unless we make him an object of our admiration. His threat of punishment gives us reasons for the relevant pro-attitudes but it does not make the demon admirable or the saucer of mud desirable. To take another case, familiar to moral philosophers, sophisticated versions of hedonism imply that we have reasons to favour, for their own sake, various other things apart from pleasure, in view of the fact that such favouring tends to be pleasure-maximizing. We are likely to be happier if we care for our friends or family for their own sake, or if we pursue knowledge in a disinterested way. But it is nonetheless only pleasure that has final value on the hedonist view. Clearly, we do not want the value analysis to rule out such hedonistic axiologies as incoherent.

### *Failed solutions*

In these examples, it would appear that there can be good reasons for pro-attitudes towards valueless objects. These reasons, however, are of the wrong kind, from the point of view of the buck-passing account: They do not bear on the value of the objects. To solve this problem, the obvious strategy would be to come up with a demarcation criterion that distinguishes the ‘right’ kind of reasons from the wrong kind. In the aforementioned paper, we go through a number of possible proposals for such a criterion. We show, however, that none of them is fully satisfactory. The proposal we find most attractive is unfortunately no exception. On that proposal, which we call the ‘dual role’-solution, reasons of the right kind are taken to play two roles: they both justify a pro-attitude and figure in its intentional content as the ground for

---

<sup>5</sup> See his “Value ... And What Follows. By Joel Kupperman,” *Philosophy* 75, (2000): 458-62.

favouring. Thus, for example, a person's courage is a right kind of reason for admiring her, because it both justifies the admiration and is the property *for* which she ought to be admired. By contrast, while the demon's threat of punishment does justify our desiring the saucer of mud for its own sake, it is a wrong kind of reason, since the threat of punishment cannot reasonably figure in the very content of such a desire as its justifying ground. We can desire the mud, say, for its texture or even for its taste, but not, it seems, for it being such that we would be punished otherwise. Likewise, the pleasure-maximizing tendency of our desire for knowledge justifies this attitude but cannot figure in its intentional content as the feature for which knowledge is being desired. This 'dual role'-solution takes care of many of the counter-examples to the buck-passing analysis, but it fails in special cases. Thus, for example, it fails in the case of the demon who demands to be admired, on pain of punishment. It does seem possible to admire someone for his determination to punish us otherwise, even though, intuitively, this reason for admiration is of the wrong kind: It does not make the demon admirable.

Olson goes through some of the proposals we have discussed (he does it in an admirably elegant and concise way) and he rejects them on the same grounds as we have done. Then he delineates his own solution. To prepare the ground for it, we first need to say something about one of the proposals both he and we reject. On that suggestion, all reasons of the wrong kind are attitude- rather than object-given: They are supplied by the properties of the pro-attitude and not by the properties of its object.<sup>6</sup> On the other hand, if it is a property of the object that provides a reason, then, and only then, we have a reason of the right kind. The proposal is attractive, since reasons of the wrong kind intuitively bear on the value of the pro-attitude, or

---

<sup>6</sup> See Parfit's distinction between 'state-given' (= attitude-given) reasons and 'object given' reasons in his "Rationality and Reasons," in Dan Egonsson, Björn Petersson, Jonas Josefsson, and Toni Rønnow-Rasmussen (eds.), *Exploring Practical Philosophy: From Action to Values* (Aldershot: Ashgate, 2001), pp. 17 – 41. It was that distinction that led us to consider this proposal. Parfit himself, however, did not use the distinction to solve problems in value analysis.

on its deontological status, while reasons of the right kind have bearing on the value of the object.

The proposal leads to difficulties, though, if one allows for proliferation in property ascriptions. As Olson succinctly puts it:

[I]t is easily seen that properties of the attitudes may be recast as properties of the objects; if, e.g., the attitude of preferring the saucer of mud has the property of preventing our suffering severe pain, then the saucer of mud has the corresponding property of being such that preferring it would prevent our suffering severe pain. (p. 299)

If the property of the attitude provides a reason for that attitude, then, surely, the corresponding property of the object also provides a reason for that attitude. This means, then, that on the proposal under consideration there is a right kind of reason for a pro-attitude whenever there is a wrong kind of reason, which makes the suggested criterion totally ineffective.<sup>7</sup>

One might try to meet this objection by insisting that it is only ‘real’ properties of objects that can provide reasons of the right kind. That the saucer of mud is “such that preferring it would prevent our suffering severe pain” is a mere ‘Cambridge property’. This is an attractive suggestion, but it does not carry all the way. Some properties of objects that intuitively provide reasons of the wrong kind do not have such Cambridge-like character. A case in point is the example in which the demon threatens us with punishment if we refuse him admiration. There is nothing Cambridge-like in his determination to punish us if we don’t comply. But still, this feature of the demon is a reason of the wrong kind: While being a strong reason for admiring him, it does not make him admirable.<sup>8</sup> Olson makes the same point using another example. As he also points out, it might not be a “good tactic” for a buck-passer “to rest the response to the WKR objection on an illiberal way of construing what is to count as a property.” (p. 299) An analysis of value that relies on a contentious theory of properties,

---

<sup>7</sup> Cf. “The Strike of the Demon”, p. 406.

<sup>8</sup> See “The Strike of Demon”, p. 407f.

which disallows Cambridge-like features, makes itself vulnerable to potential metaphysical criticisms. Anyway, even if one is prepared to pay this cost, the proposal is unable to deal with some counter-examples, as we have just seen.

### *Olson's solution*

Olson's own solution is to impose a more severe restriction on the content of a statement supplying a reason for an attitude. He refers to such statements as "JE-clauses", with "JE" standing for "justificatory explanation". In his interpretation of the notion of a reason, he follows John Broome: To give a reason is to provide an explanation. A (normative) reason to  $\phi$  explains why one *ought* to  $\phi$ .<sup>9</sup> In other words, such reasons constitute justificatory explanations. Examples of JE-clauses are: "the demon will punish us unless we desire the saucer of mud," "the demon will punish us unless we admire him", or, to take the hedonist type of example, "caring for your friend for his own sake is pleasure-maximizing". These are all examples of clauses that state reasons for pro-attitudes that intuitively are of the wrong kind. Examples of reasons of the right kind are given by clauses such as "she is courageous" (a reason for admiration), "this experience is pleasant" (a reason for desiring it), or, say, "this wilderness has never been touched by a human hand" (a reason for preservation and care). The question is, then, what distinguishes the latter group of reason clauses from the former.

Here is Olson's solution:

Since [as we have seen] properties of the attitude (*A*-properties) are easily translatable into properties of the objects (*O*-properties), it is insufficient to demand that the JE-clause must contain only *O*-properties and no *A*-properties. We need to make the stricter demand that JE-clauses must not be, as I shall say, *A-referential*. That is, JE-clauses must not contain any reference whatsoever to properties of the attitude in question, whether in the guise of properties of the attitude or, more limitedly, of the object. (p. 299)

It is not very easy to understand precisely what Olson has in mind. This talk of "properties of the attitude ... in the guise of properties of the attitude or ... of the object" is rather

---

<sup>9</sup> See Broome, "Reasons", in R. Jay Wallace *et al.* (eds.), *Reasons and Value: Themes from the Moral Philosophy of Joseph Raz*, OUP 2004, pp. 28-55.

confusing. But here is what we think he is after. Consider the clause “the demon will punish us unless we admire him”. On a natural reading, this statement ascribes a certain property to the demon, but it could also be read as an ascription of a property to the admiration of the demon, viz., that this attitude would shield us from his punishment. When one considers the different JE-clauses above that state reasons of the wrong kind, one notices that all of them explicitly mention the attitude for which they supply a reason. Thus, each can be read as ascribing a property to the attitude in question (even though they can also be read as ascribing a property to the object). It is this, on Olson’s view, that makes them statements of reasons of the wrong kind. The JE-clauses in the second group, on the other hand, do not refer to the attitudes they supply reasons for. Consequently, they cannot be seen as ascriptions of properties to these attitudes. This is what makes them statements of the right kind of reasons.

### *Criticism*

This solution is attractive and natural, since in the case of the wrong kind of reasons it is always something about the pro-attitude itself that explains why the attitude ought to be taken. Consequently, by banning references to the relevant attitudes in the reason-statements, we seem to keep out all statements that formulate reasons of the wrong kind. The question is, however, whether the ban on references to attitudes and their properties does not go too far. Aren’t there some reasons of the *right* kind that will be hit by this prohibition? Won’t we be throwing out some of the babies with the bath water?

Unfortunately, it seems that we will. In some cases, a property of an object that supplies a reason of a right kind might well relate in some way to the pro-attitudes for which it provides a reason. Think, for example, of a person who is admirable because of certain character traits. Such a person can be even more admirable if he doesn’t care about whether he is being admired. He is ‘above’ concerns of this kind. Needless to say, indifference to being admired cannot itself make a person admirable, but it can make him *more* admirable if he otherwise

deserves admiration. Thus, indifference to admiration is a property that can provide a (further) reason of the right kind for admiring a person, even though that property involves a relativisation to the very pro-attitude for which it provides a reason.

Consider another example, due to John Broome (personal communication): Suppose a person has a disposition to respond with love to love. We might well think that this character trait makes him lovable (with “lovability” understood as a value term). Note that this dispositional feature might make loving such a person instrumentally valuable (*if* being loved is of value), but it might also, not implausibly, make him a fitting object of love for his own sake. So, again, we have an example of a property that makes a person finally valuable, even though the ascription of this property in a JE-clause makes a reference to the pro-attitudes for which it provides a reason. Obviously, examples of such attitude-referring reasons of the right kind can be easily multiplied, once one gets the idea of how to look for them in the first place.

One might try to object to these counterexamples in the following way: On the proposal we consider, a JE-clause states a reason of the right kind if and only if it does not refer to the very pro-attitude for which it provides a reason. But in each of our examples of attitude-referring reasons of the right kind, the JE-clause does not make reference to any particular token of the pro-attitude in question, nor does it refer to any limited set of such tokens. Instead, it refers to an attitude *type* (love, admiration). Does it show that the counterexamples we use are illegitimate?

We do not think so. Consider examples of JE-clauses that state reasons of the *wrong* kind. For a sophisticated hedonist, the fact that, say, “caring for knowledge, for its own sake, is pleasure-maximizing” provides a reason for having this attitude toward knowledge. In the statement of the reason, though, the reference is made not to a particular token of the attitude but to an attitude type. We have a reason to care for knowledge for its own sake, because, on the hedonist view, attitudes of this type have beneficial consequences (= a reason of the



wrong kind). In this respect, then, there seems to be no difference between reasons of the wrong kind and attitude-referring reasons of the right kind. Olson's criterion would be much too weak if it only banned references to attitude tokens. The counter-examples we have provided cannot therefore be dismissed in this way.

One might instead try to respond to our counter-examples by modifying Olson's proposal. On that proposal, a JE-clause states a reason of the right kind if and *only if* it does not refer to the pro-attitude for which it provides a reason. As we have seen, this is too restrictive. Perhaps, however, one might weaken this requirement somewhat. What if we only require that the right reason for an attitude should at least be *derived* from some reason for that attitude that is expressible without any reference to the attitude in question? Thus, for example, one might suggest that a person's indifference to being admired shows that he is modest and that the former property is a reason for admiration only insofar as *modesty* is such a reason. Similarly, it might be suggested that a person's disposition to respond with love to love shows that he has ability to love and that the former property is a reason for love only insofar as the latter property is such a reason. These suggestions concerning our two counterexamples have been made by an anonymous referee.

This modified version of Olson's proposal is challenging but nonetheless unconvincing:

(i) In some cases, it seems, the reference to the relevant attitude will not really be avoided by a move to the underlying reason. It will only be hidden from sight. This is what we believe happens when we move from indifference-to-admiration to modesty as the underlying reason. After all, being modest seems to consist, at least in part, in being indifferent to praise and admiration.

(ii) In other cases, the move to an underlying reason may instead be unjustified. The referee's suggestion concerning the disposition to respond with love to love can in fact serve as an example. While such a disposition might well make one a lovable person, the same need

not apply to a mere ability to love. The latter disposition is probably too unexceptionable to provide a reason for the relevant pro-attitude. To be sure, this particular example is not really important. Perhaps a better reduction to an underlying reason can be found in this case. What is important, however, is that, in general, there is simply no guarantee that for every right kind of reason involving a reference to the relevant attitude it will always be possible to find an underlying reason that does not involve such a reference.

To conclude, therefore, we do not think that this modification of Olson's proposal will be successful.

### *Circular buck-passing*

Does it mean, then, that the WKR-problem necessitates giving up the buck-passing account? Not quite. In the first place, there is still hope that the problem can be solved, even though Olson's proposal fails to accomplish this goal. Secondly, even if the problem would turn out to be unsolvable, we could continue to hold on to the buck-passing account, *if* we are prepared to allow philosophical analyses that involve some amount of circularity. After all, we can always define reasons of the right kind as precisely those reasons for pro-attitudes that invoke the value-making properties of their objects. Thereby, we make the analysis of value in terms of the right kind of reasons explicitly circular, but such circles might be acceptable: Circular analyses still allow us to exhibit structural connections between central concepts (value, reason, pro-attitude). Thereby, they can provide relevant information to those who have the concepts but are not clear about their mutual relationships.

This approach to the buck-passing account is taken by David Wiggins.<sup>10</sup> Wiggins adheres to a sentimentalist version of buck-passing: He assumes that the pro-attitudes that are called for by valuable objects are affective in nature. The circularity Wiggins has in mind does not arise because of the WKR-problem. Instead, it concerns the presence of evaluations in

---

<sup>10</sup> See "A Sensible Subjectivism?," essay V in D. Wiggins, *Needs, Values, Truth: Essays in the philosophy of value* (Oxford: Blackwell, 1987), p. 189.

sentiments: According to him, the evaluative element is essential to affective attitudes. Consequently, understanding sentiments presupposes understanding the concept of evaluation.

Now, we owe to I. L. Humberstone a useful distinction between two kinds of analytical circularities.<sup>11</sup> There are relatively benign cases, in which the analysandum appears in the analysans only within a ‘protective’ linguistic context, say, within a scope of a doxastic operator. These can be contrasted with ‘inferential circularities’, in which the analysandum appears unprotected in the analysans. The latter type of circularity is more serious according to Humberstone, since it makes it impossible to determine whether the analysans applies to a given case as long as one hasn’t yet determined whether the analysandum is applicable.<sup>12</sup> As is easily seen, Wiggins’ circularity is not of the inferential kind. When we analyse the value of an object in terms of fitting sentiments and then go on to analyse sentiments as attitudes involving value ascriptions, the context in which the value concept appears in the analysans is protective: We can know whether someone ascribes value to an object without knowing whether that object is valuable. But if reasons of the right kind are defined as reasons that invoke value-making properties, the buck-passing account becomes inferentially circular. Thus, this circle is more serious than the one entertained by Wiggins.

Still, it seems to us that even inferential circularities can be tolerated in philosophical analyses of central concepts. The reason for providing an analysis in such cases is not so much the need of criteria for a correct application of the concepts under consideration, but rather the importance of gaining a deeper understanding of the conceptual framework. Tracing mutual connections between the concepts involved can be of help in this task. To this extent, then, inferential circularity need not vitiate the main purposes of the analytical enterprise.

---

<sup>11</sup> See his “Two Types of Circles”, *Philosophical and Phenomenological Research* 57 (1997), pp. 249-80.

<sup>12</sup> For a detailed critical discussion of Humberstone’s distinction, see Rosanna Keefe, “When Does Circularity Matter?”, *Proceedings of Aristotelian Society* 102 (2002), pp. 253-270.