

Wrong Kind of Reasons and Consequences

RICHARD ROWLAND

University of Reading

In a recent issue of *Utilitas* Gerald Lang provided an appealing new solution to the Wrong Kind of Reason problem for the buck-passing account of value. In subsequent issues Jonas Olson and John Brunero have provided objections to Lang's solution. I argue that Brunero's objection is not a problem for Lang's solution, and that a revised version of Lang's solution avoids Olson's objections. I conclude that we can solve the Wrong Kind of Reason problem, and that the wrong kind of reasons for pro-attitudes are reasons that would not still be reasons for pro-attitudes if it were not for the additional consequences of having those pro-attitudes.

T. M. Scanlon's buck-passing account of goodness and value (BPA) holds that good *simpliciter* and final value can be analysed in the following way:

*BPA: X is good simpliciter or of final value if and only if X has properties that give us reasons to have a certain pro-attitude towards X (such as to admire X or to desire X for its own sake).*¹

The most famous objection to the BPA is that it over-generates instances of things that are good *simpliciter* and of final value. According to this objection there are things that there are reasons for us to have a pro-attitude towards that are not good *simpliciter* or of final value.² Consider the following two examples:

Desire a Saucer of Mud: An evil demon will severely punish us if we do not desire a saucer of mud for its own sake. So, there is a reason for

¹ See T. M. Scanlon, *What We Owe to Each Other* (Cambridge, Mass., 1998), pp. 95–7, and Gerald Lang, 'The Right Kind of Solution to the Wrong Kind of Reason Problem', *Utilitas* 20 (2008), pp. 472–89. As Stratton-Lake and Hooker point out, 'give us reasons' should be understood as 'give everyone reasons'. I use 'us' instead of 'everyone' throughout this article because those I discuss use 'us' rather than 'everyone'; see Philip Stratton-Lake and Brad Hooker, 'Scanlon versus Moore on Goodness', *Metaethics After Moore*, ed. Mark Timmons (Oxford, 2006), pp. 149–68, at 152–3.

² Hereafter I will use 'goodness' and 'value' to refer to goodness *simpliciter* and final value. Something is good *simpliciter* if it is good for its own sake, good full stop, or non-elliptically good. This type of goodness is the type of goodness that Peter Geach and Judith Jarvis Thomson deny exists; see P. T. Geach, 'Good and Evil', *Analysis* 17 (1956), pp. 33–42, and Judith Jarvis Thomson, *Normativity* (Ithaca, 2008), pp. 1–17. Something is of final value if it is of value for its own sake, if it is of non-instrumental non-attributive value; see Christine Korsgaard, 'Two Distinctions in Goodness', *Philosophical Review* 92 (1983), pp. 169–95.

us to desire a saucer of mud for its own sake. But the saucer of mud is not good or of value.

Admire the Demon: An evil demon will severely punish us if we do not admire him. So, there is a reason for us to admire the evil demon. But the evil demon is not good or of value.³

These examples are examples in which a saucer of mud and an evil demon satisfy the conditions on the right-hand side of the BPA's 'if and only if', but are not good or of value.

Examples like *Desire a Saucer of Mud* and *Admire the Demon* lead to the Wrong Kind of Reason problem for the BPA. The Wrong Kind of Reason problem is the problem of distinguishing the *right kind of reasons* from the *wrong kind of reasons*. The right kind of reasons are given by properties that satisfy the conditions on the right-hand side of the BPA's 'if and only if', and are properties of things that are good or of value. The wrong kind of reasons are given by properties that satisfy the conditions on the right-hand side of the BPA's 'if and only if', but are properties of things that are not good or of value. Proponents of the BPA must distinguish between the right kind of reasons and the wrong kind of reasons in a way that does not rely on the concepts of goodness and value that are being analysed (in the BPA) and is not ad hoc or artificial.

LANG'S SOLUTION TO THE WRONG KIND OF REASON PROBLEM

In a recent article in this journal, Gerald Lang provides a compelling solution to the Wrong Kind of Reason problem. According to Lang's solution, wrong kind of reasons for us to have pro-attitudes are reasons that would not be reasons for us to have these pro-attitudes if having these pro-attitudes would not benefit us. And having a pro-attitude benefits someone if they would be better off having that attitude than not having that attitude.⁴ More formally, according to Lang, the BPA should be modified in the following way to solve the Wrong Kind of Reason problem:

BPA Benefits: X is good *simpliciter* or of final value if and only if X has properties that give us reasons to have a pro-attitude towards X and those properties of X would still give us reasons to have a pro-attitude towards X in the absence of the benefits to us of having that attitude towards X .⁵

³ See Wlodek Rabinowicz and Toni Rønnow-Rasmussen, 'The Strike of the Demon: On Fitting Pro-Attitudes and Value', *Ethics* 114 (2004), pp. 391–423, at 405–7.

⁴ See Lang, 'Right Kind of Solution', p. 484, and Jonas Olson, 'The Wrong Kind of Solution to the Wrong Kind of Reason Problem', *Utilitas* 21 (2009), pp. 225–32, at 226.

⁵ See Lang, 'Right Kind of Solution', p. 484.

On Lang's solution the reason for us to desire a saucer of mud for its own sake in *Desire a Saucer of Mud* comes out as a wrong kind of reason because there would be no reason for us to desire a saucer of mud for its own sake if we would not benefit from so desiring a saucer of mud. Similarly, the reason for us to admire the demon in *Admire the Demon* comes out as a wrong kind of reason on Lang's solution because there would be no reason for us to admire the demon if we would not benefit from admiring it.

In the rest of this article I discuss John Brunero's and Jonas Olson's recent objections to Lang's solution to the Wrong Kind of Reason problem. I argue that Brunero's objections do not undermine Lang's solution and that Lang's solution can be modified so as to avoid Olson's objections.

BRUNERO'S RESPONSE TO LANG

In his reply to Lang, John Brunero claims that Lang's solution to the Wrong Kind of Reason problem is incompatible with two types of consequentialism.⁶ First, Brunero claims that Lang's solution is incompatible with a consequentialist view, according to which there are reasons to have pro-attitudes towards things that are *not* of final value just because of the benefits of having these pro-attitudes. Brunero calls this view indirect consequentialism.

Consider a version of this type of consequentialism according to which pleasure is the only thing that is of final value. Call this view hedonic indirect consequentialism. According to hedonic indirect consequentialism:

- (1) Pleasure is the only thing that is of final value.

And

- (2) There are reasons to have a pro-attitude, such as to aim to have friends, but there would be no reason to aim to have friends if there were no good consequences to aiming to have friends. There is a reason for us to aim to have friends only if aiming to have friends will produce more pleasure.

According to Brunero, (3) follows from (1) and (2):

- (3) Pleasure has features that give us reasons to have a pro-attitude, namely to aim to have friends, but these reasons to aim to have

⁶ Brunero might hold that these two types of consequentialism, which he calls direct and indirect, cover all the possible consequentialist views. However, this is not clear in general, and is far less clear given the way Brunero describes these views.

friends would not persist in the absence of the consequences of aiming to have friends.

But, according to Lang's solution to the Wrong Kind of Reason problem,

- (4) *X* is of final value if and only if *X* has features that give us reasons to have a pro-attitude, and *X* would give us reason to have a pro-attitude if there were no benefits to having this pro-attitude.

But (5) follows from (3) and (4):

- (5) Pleasure is not of final value.

And (6) follows from (1) and (5):

- (6) Pleasure is both of final value and not of final value.

So, according to Brunero, Lang's solution to the Wrong Kind of Reason problem is incompatible with indirect consequentialism because the combination of the two views yields a contradiction.⁷

But (4) importantly misrepresents Lang's solution to the Wrong Kind of Reason problem. Lang's solution, and the BPA, does not hold that *X* is of final value if and only if *X* has features that give us reasons to have a pro-attitude, but rather that *X* is of final value if and only if *X* has features that give us reasons to have a pro-attitude *towards X*. So (4) should be modified to (4'):

- (4') *X* is of final value if and only if *X* has features that give us reasons to have a pro-attitude *towards X*, and *X* would give us reasons to have a pro-attitude if there were no benefits to having this pro-attitude.

But (3) and (4') do not entail that pleasure is not of final value, since according to hedonic indirect consequentialism pleasure gives us reasons to have a pro-attitude *towards friendship* and not towards pleasure. So, Lang's solution is not incompatible with indirect consequentialism.⁸

⁷ See John Brunero, 'Consequentialism and the Wrong Kind of Reasons: A Reply to Lang', *Utilitas* 22 (2010), pp. 351–9, at 357. Brunero detailed and assessed this case in this way in private correspondence. I detail this case and not the analogous case discussed in Brunero's article because it is so easy to mistake Brunero's argument with regard to the case in his article for an argument that is a non-starter because it equivocates between final and instrumental value.

⁸ Even if aiming to have friends would be having a pro-attitude towards pleasure in this case, (3) and (4') would not entail that pleasure is not of final value. Pleasure not only gives us reasons to aim to have friends in this case; pleasure gives us reasons to have other pro-attitudes such as to desire it for its own sake and to wish that others had it. And pleasure would still give us reasons to desire it for its own sake and to wish that others had it even if there were no benefits to our desiring it for its own sake and wishing that others had it.

Brunero claims that Lang's solution is incompatible with another type of consequentialism, which Brunero calls direct consequentialism. According to this type of consequentialism, we should aim to promote value. But there would be no reason for us to aim to promote value if aiming to promote value did not have the beneficial consequence of making it more likely that we will promote value.

Consider a version of this type of consequentialism according to which pleasure is the only thing that is of final value. Call this view hedonic direct consequentialism. According to hedonic direct consequentialism:

(1*) Pleasure is the only thing that is of final value.

And

(2*) There are reasons to have a pro-attitude towards pleasure, specifically to aim to promote pleasure. But there would be no reason to aim to promote pleasure if there were no benefits to aiming to promote pleasure. Or rather, there are reasons to aim to promote pleasure only if aiming to promote pleasure has the beneficial consequence of making it more likely that we will promote pleasure.

But, according to Lang's solution to the Wrong Kind of Reason problem,

(3*) *X* is of final value if and only if *X* has features that give us reasons to have a pro-attitude towards *X*, and *X* would give us reasons to have a pro-attitude if there were no benefits to having this pro-attitude.

And according to Brunero (4*) follows from (2*) and (3*):

(4*) Pleasure is not of final value.

And (5*) follows from (1*) and (4*):

(5*) Pleasure is both of final value and not of final value.

So, Lang's solution to the Wrong Kind of Reason problem is incompatible with direct consequentialism because the combination of the two views yields a contradiction.⁹

If pleasure only gave us reasons to aim to promote pleasure, it would follow from the combination of Lang's solution and part of hedonic direct consequentialism – namely (2*) – that pleasure is not of final value. But pleasure does not *only* give us reasons to aim to promote pleasure. Pleasure gives us reasons to have other pro-attitudes such as to desire it for its own sake and to wish that we could promote

⁹ See Brunero, 'Consequentialism', p. 357.

it. And there would be reasons for us to desire pleasure and wish we could promote pleasure even if desiring it and wishing that we could promote it had no beneficial consequences. Whatever is of final value, that which is of final value gives us reasons to desire it for its own sake or to desire that it is preserved for its own sake regardless of the beneficial consequences of our desiring these things.

Lang's solution and (2*) would entail that pleasure is not good or of value in this case only if Lang's solution entailed the following claim:

If X gives us reasons to have pro-attitudes a , b , c , and d , but the reasons to have pro-attitudes a and b would not persist in the absence of the benefits of having pro-attitudes a and b – even though the reasons to have c and d would – then X is not of final value.

But Lang's solution does not entail this claim. So, the combination of Lang's solution and (2*) does not entail that pleasure is not of final value – (4*) does not follow from (2*) and (3*). Thus, Lang's solution is not incompatible with this type of consequentialism either.

CONSEQUENTIALISM ABOUT ABSOLUTELY EVERYTHING

There is a worry that lies behind both of Brunero's objections. Some consequentialists, consequentialists about absolutely everything, hold a consequentialist account of reasons for pro-attitudes.¹⁰ These consequentialists about everything hold that there is a reason to have a pro-attitude only if having that pro-attitude has good consequences; there is a reason to ϕ only if ϕ -ing has good consequences.¹¹ So, consequentialists about everything will hold that there are reasons to desire something for its own sake only if there are good consequences to desiring that thing for its own sake.

Notice that consequentialism about everything is far less plausible than the view, which Brunero describes, according to which there is a reason to aim to ϕ only if there are good consequences to aiming to ϕ . It might seem that there is no reason for me to *aim* to fly like a bird when I cannot fly like a bird and my aiming to fly like a bird has no good consequences. But there is of course a reason for me to *wish* that I could fly like a bird even though I cannot fly like a bird and even if my wishing that I could fly like a bird has no good consequences. Similarly, there is a reason for us all to desire to avoid excruciating pain, even if we are being tortured and cannot avoid excruciating pain, and our desiring to avoid such pain will not enable us to avoid this pain.

¹⁰ This view is often called global consequentialism.

¹¹ Brunero made clear that this was the worry behind his objections in private correspondence.

However, regardless of whether consequentialism about absolutely everything is plausible, and whether a consequentialist constraint on *all* reasons for pro-attitudes is plausible, it is not a problem for Lang's solution that it is not compatible with such a consequentialist constraint on reasons for pro-attitudes. The BPA is an account of good *simpliciter* and final value. As such it must be compatible with all substantive views about what things are good *simpliciter* or of final value.¹² The BPA must not take sides on whether pleasure is the only thing that is good, for instance. But the BPA need not be compatible with all accounts of normative reasons or views about when there are normative reasons. And a consequentialist account of, or constraint on, reasons for pro-attitudes is not a consequentialist account of what things are good or of value but an account of normative reasons for pro-attitudes. So it is not a problem for the BPA if it is not compatible with this consequentialist account of reasons for pro-attitudes. And so it is not a problem for Lang's solution that it is not compatible with a consequentialist account of reasons for pro-attitudes.

Furthermore, it seems that this type of consequentialism about reasons for pro-attitudes was always going to be in conflict with the BPA. This type of consequentialism about absolutely everything involves an account of the relationship between reasons and the good; it holds that reasons are defined in terms of, or derive from, the good. But the BPA also explicitly provides an account of the relationship between reasons and the good; according to the BPA, the good is defined in terms of reasons. Consequentialism about everything provides an account of reasons in terms of the good, and the BPA provides an account of goodness in terms of reasons. And because these two views of the relationship between reasons and the good are incompatible, consequentialism about everything was always going to be incompatible with the BPA, and not only Lang's solution to the Wrong Kind of Reason problem for the BPA.

Lars Samuelsson understands the worry behind Brunero's objections differently. According to Samuelsson, if having an

attitude would not help to bring about value, there is no reason why a consequentialist should insist that we have reason to have that attitude.

Hence, any version of [the BPA] which takes values to correspond to reasons for *attitudes* . . . is incompatible with at least many forms of consequentialism, and should therefore be rejected as a formal account of value.¹³

¹² See Rabinowicz and Rønnow-Rasmussen, 'Strike of the Demon', p. 403, and Lang, 'Right Kind of Solution', p. 474.

¹³ Lars Samuelsson, 'The Right Version of "the Right Kind of Solution to the Wrong Kind of Reason Problem"', *Utilitas* (2013), doi:10.1017/S095382081200057X.

Samuelsson concludes that BPAs in terms of reasons for pro-attitudes are incompatible with consequentialism.

I presume that by 'a consequentialist' Samuelsson has in mind someone who holds the position that is normally referred to by this name in normative ethics, that is, someone who holds the position that an act's rightness or wrongness depends on its consequences. But this sort of consequentialist need not accept that we have reasons to have pro-attitudes towards value even when there are no good consequences to doing so because this sort of consequentialist *qua* this sort of consequentialist only holds a view about rightness and wrongness and not a view about reasons for pro-attitudes. Consequentialism about reasons for action and reasons for attitudes are views distinct from consequentialism in normative ethics that, as I've been arguing, the BPA need not be compatible with, and was always going to be incompatible with.

Samuelsson, however, seems to think that consequentialism about reasons for attitudes is a more ordinary consequentialist view, or a view more closely linked to consequentialism in normative ethics, than consequentialism about reasons for action. And Samuelsson claims that any BPA should be made in terms of reasons for *pro-responses* rather than reasons for *pro-attitudes* in order to make it compatible with a consequentialist view of reasons for pro-attitudes.¹⁴

It might be that the BPA should be made in terms of reasons for pro-responses rather than reasons for pro-attitudes. Indeed Ewing, who endorsed the BPA at least thirty-five years before Scanlon, understood 'pro-attitudes' to include actions such as pursuit, and not just (mental) attitudes such as desire and admiration.¹⁵ However, a BPA in terms of *only* reasons for pro-actions is a non-starter. And in order for the BPA to be compatible with a consequentialist view of reasons for attitudes the BPA would have to be made in terms of only reasons for pro-actions and not reasons for pro-actions or pro-attitudes.

To see that a BPA in terms of reasons for pro-actions alone is a non-starter, imagine an isolated world in which there is a being that is in a pleasurable state but is incapable of action. Hedonists will hold that there is value in this world. But there are certainly no reasons to perform actions in response to the value in this world; we cannot preserve or promote the pleasure in this world, since the world is isolated. And there is no one in that world who can promote or preserve the pleasure either. The BPA must be able to account for the

¹⁴ L. Samuelsson, "The Right Version of 'the Right Kind of Solution to the Wrong Kind of Reason Problem'", *Utilitas* 25 (2013), pp. 383–404.

¹⁵ See A. C. Ewing, *The Definition of Good* (London, 1947), esp. p. 149, A. C. Ewing, *Second Thoughts in Moral Philosophy* (London, 1959), pp. 86–90, and Rabinowicz and Rønnow-Rasmussen, 'Strike of the Demon', pp. 395–7.

possibility that this being's pleasure is of value because the BPA must be compatible with hedonism in order to be an adequate formal account of value.¹⁶ This is easily done by holding that there are reasons for this being to hope that her pleasure continues, and for us to be glad that she is in a state of pleasure and to hope that it continues. But according to a BPA in terms of only reasons for actions there is no value in this world, since there are no reasons for action in this world.¹⁷ So, we should reject a BPA in terms of only reasons for pro-actions.

So, Brunero's objection to Lang's solution to the Wrong Kind of Reason problem does not undermine Lang's solution. And we should not modify the BPA or Lang's solution in order to make it compatible with consequentialism about absolutely everything or a consequentialist account of or constraint on reasons for pro-attitudes.

OLSON'S RESPONSE TO LANG

In his reply to Lang, Jonas Olson raises several different objections to Lang's solution to the Wrong Kind of Reason problem. First, Olson imagines that an evil demon will punish *others* and not us if we do not admire him. It seems that there is a reason for us to admire the demon in this case. This is a problem for Lang's solution because according to his solution a reason is of the right kind if and only if it would still be a reason for us to have a pro-attitude in the absence of the benefits *to us* of having that pro-attitude. So Lang's solution entails that a demon that will punish *others* – rather than *us* – if we do not admire him is of final value.

Second, Olson imagines that unless we admire an evil demon, that demon will create a number of individuals whom he will make suffer greatly. But if we admire this demon, no one would benefit from our

¹⁶ See above, n. 12.

¹⁷ Samuelsson might claim that we should endorse a BPA in terms of counterfactual reasons for action. In this case, the being's pleasure is of value because it has properties that would provide reasons to promote it if anyone could promote it; see Jussi Suikkanen, 'Reasons and Value: In Defence of the Buck-Passing Account', *Ethical Theory and Moral Practice* 7 (2004), pp. 513–35, at 532–3. However, there are several problems with a counterfactual BPA. The first is that there could be value even if no one could possibly have pro-actions in response to it. Suppose that there were a being which cannot act, is in pleasure, and to which no one could ever respond with a pro-action. This being's pleasure might be of value even though no one could possibly have a pro-action in response to this value, and so it seems like the counterfactual BPA entails that, *contra* hedonism, this being's pleasure is not of value. Second, imagine a creature that is in an isolated world in a state of pleasure, but would no longer be in pleasure if there were beings that could respond to reasons or perform actions around. In this case the counterfactual BPA fails also because this creature's pleasure is of value – according to hedonism, for example – but its pleasure would not provide reasons to perform pro-actions in response to it if there were beings that could perform pro-actions in response to it, because if there were such beings, then it would not be in a state of pleasure.

admiring him, since there would be no one who did not suffer greatly, and we cannot benefit individuals who do not and will not exist. So, according to Lang's solution, if the demon will create a number of individuals whom he will make suffer greatly unless we admire him, then the demon is good or of value because our admiring the demon would not benefit anyone. The consequence of our admiring the demon would be that the demon did not create a number of suffering individuals; no one would be benefited.¹⁸

Third, we can imagine cases in which an evil demon will make the world less valuable in a way that does not harm anyone unless we admire him. Suppose that unless we admire him a demon will make the world less beautiful than it is now, and the demon's making the world less beautiful will not harm anyone. (The demon will, for instance, erase everyone's memories of the world being as beautiful as it is now.) In this case it seems that there is a reason for us to admire the demon because the demon will make the world a less beautiful place if we do not. But there are no benefits to our admiring the demon because we will not be harmed if the demon does make the world a less beautiful place.¹⁹ And, according to Lang's solution, X is of final value if and only if X would still give us reasons to have a pro-attitude towards X regardless of the benefits to us of our having a pro-attitude towards X . So, according to Lang's solution, an evil demon who will make the world less beautiful – in a way that will not harm us – if we do not admire him is good or of value.

But Lang's solution to the Wrong Kind of Reason problem can be revised in a way that avoids Olson's three objections. Instead of distinguishing between wrong kind of reasons and right kind of reasons in terms of counterfactual benefits, Lang's solution can be amended so as to distinguish between wrong kind of reasons and right kind of reasons in terms of counterfactual additional consequences. According to this revised version of Lang's solution:

BPA Additional Consequences: X is good *simpliciter* or of final value if and only if X has properties that give us reasons to have a pro-attitude towards X regardless of the additional consequences of our having a pro-attitude towards X .

An additional consequence of P is a logical consequence of P that would not have obtained were it not for P . When a demon will punish us if we do not admire him our avoiding punishment from the demon is an additional consequence of our admiring the demon because if we did

¹⁸ Olson, 'Wrong Kind of Solution', pp. 226–7.

¹⁹ Olson, 'Wrong Kind of Solution', p. 227.

not admire the demon, we would not have avoided punishment. $2 + 2 = 4$ might be a logical consequence of our admiring the demon. But the fact that $2 + 2 = 4$ is not an *additional* consequence of our admiring the demon, since if we had not admired the demon it would still have been the case that $2 + 2 = 4$.

BPA Additional Consequences avoids Olson's objections. It avoids Olson's first objection because one additional consequence of our admiring a demon that will punish others if we do not admire him is that others are not punished. If it were not the case that our admiring such a demon would have the additional consequence that others are not punished, there would be no reason for us to admire him.²⁰

BPA Additional Consequences also avoids Olson's second objection. When an evil demon will bring new agents into the world that will suffer greatly unless we admire him there is a reason for us to admire the demon. But there would be no such reason if it weren't for one of the additional consequences of our admiring him. One of the additional consequences of our admiring the demon in this case is that the demon will not bring agents into the world who will suffer greatly.

And *BPA Additional Consequences* avoids Olson's third objection because there would be no reason for us to admire a demon that will make the world less beautiful unless we admire it if it were not for the additional consequences of our admiring such a demon. One of the additional consequences of our admiring such a demon would be that the world is not made less beautiful; if this were not one of the additional consequences of our admiring the demon, there would be no reason for us to admire it.

CONCLUSION

Contra Brunero, Lang's solution to the Wrong Kind of Reason problem does not conflict with what Brunero calls indirect or direct consequentialism. Although Lang's solution does conflict with a consequentialist account of and constraint on reasons for pro-attitudes, the buck-passing account of value need not be compatible with all accounts of and constraints on reasons for pro-attitudes. And a revised version of Lang's solution that distinguishes between wrong kind of reasons and right kind of reasons in terms of counterfactual additional consequences rather than in terms of counterfactual benefits

²⁰ Although it might be that the wrong kind of reasons to have pro-attitudes are not reasons to have pro-attitudes at all. See for instance Jonathan Way, 'Transmission and the Wrong Kind of Reason', *Ethics* 122 (2012), pp. 489–515.

circumvents Olson's objections to Lang's solution. So, the Wrong Kind of Reason problem can be solved. Right kind of reasons for pro-attitudes are reasons that would still be reasons for pro-attitudes regardless of the additional consequences of having those pro-attitudes.²¹

r.a.rowland@pgr.reading.ac.uk

²¹ I would like to thank John Brunero, Brad Hooker and Philip Stratton-Lake for comments on and discussions of previous drafts of this article. An Arts and Humanities Research Council studentship funded the writing of this article.