Fittingness First[*]

Conor McHugh and Jonathan Way

Abstract

According to the *fitting-attitudes* account of value, for X to be good is for it to be

fitting to value X. But what is it for an attitude to be fitting? A popular recent

view is that it is for there to be sufficient *reason* for the attitude. In this paper we

argue that proponents of the fitting-attitudes account should reject this view and

instead take fittingness as basic. In this way they avoid the notorious 'wrong kind

of reason' problem, and can offer attractive accounts of reasons and good

reasoning in terms of fittingness.

According to the *fitting-attitudes* account of value, what it is for X to be good is for it to be

fitting to value X. Famously proposed by Brentano and Ewing, the fitting-attitudes account

takes goodness to be similar to properties like that of being admirable, fearful, or amusing.[1]

These properties are plausibly understood as a matter of the fittingness or appropriateness of

a certain human response: admiration, fear, or amusement. Similarly, on the fitting-attitudes

view, goodness, or being valu*able*, is a matter of the fittingness of the response of *valuing*.

The fitting-attitudes account is often thought to constitute an attractive middle-ground

between subjectivist views on which something is good just if we value it and Moorean views

on which goodness is a basic, non-natural – and, so critics say, mysterious – property. And while the advantages of the fitting-attitudes account over Moorean views should not be overstated – the metaphysical questions that arise about Moorean goodness seem also to arise about fittingness – the connection between value and human valuing responses does seem an attractive feature of the fitting-attitudes account.

Nonetheless, and even putting aside the question of whether fittingness is a 'non-natural' property, there is an important question about how fittingness is to be understood. A popular recent suggestion is that fittingness should be understood in terms of reasons – roughly, that what it is for it to be fitting to value X is for there to be sufficient reason to value X. This idea is especially natural for those sympathetic to the recently influential *reasons-first* approach to normativity. On this approach, reasons are the basic normative unit, and the rest of the normative and evaluative domain can be understood in terms of reasons.[2] If we put this view together with the fitting-attitudes account, we reach the *buck-passing* account of value: what it is for X to be good is for there to be sufficient reason to value X.

The aim of this paper is to argue that proponents of the fitting-attitudes account should not take this path. Instead, they should take fittingness as basic. Doing so results in a view which has all the attractions of the buck-passing view, and more, but avoids a serious problem facing the buck-passer. In making this case, we also show how reasons (and thus oughts) can be understood in terms of fittingness. The paper thus makes an initial case for taking *fittingness first*: for seeing fittingness as the basic normative property from which the rest of the normative and evaluative domain is constructed.

The buck-passing view is sometimes understood as an attempt at conceptual analysis. However, we will understand it as a metaphysical account of the property of goodness: an account of what it is for something to be good.[3] Similarly, our fittingness-first view says that fittingness is metaphysically first, not that it is conceptually or epistemically prior to

goodness or reasons. Note also that our claim is that fittingness is first relative to the rest of the normative and evaluative domain. This leaves open the possibility – about which we are neutral – that fittingness may in turn be reducible to the natural or descriptive.

One might ask why we should expect that either reasons or fittingness comes 'first'. Why expect that there will be some basic normative or evaluative property from which the rest of the normative and evaluative domain is constructed? We do not assume from the outset that this is so.[4] However, we take it that this hypothesis has significant attractions. In general, accounts of a property F in terms of another property G promise straightforward explanations of features of F and of connections between F and G. We will discuss several examples of this in what follows. And since the normative and evaluative domain seems highly interconnected, the hypothesis that there is one basic normative property has significant explanatory promise. Whether this hypothesis can ultimately be sustained is a question this paper aims to contribute towards answering.

The paper has three main sections. The first sets out the central attractions of the buck-passing account and the central problem with it. The second outlines our alternative fittingness-first view and explains how this view retains the attractions of the buck-passing view while avoiding the problem. The third section considers objections to our view.

1. The Buck-Passing Account of Value

According to the buck-passing account of value:

What it is for X to be good is for there to be sufficient reason for anyone to value X.[5]

'Value' is used here as a cover-term for a variety of pro-responses. Examples of such responses include desire, choice, wish, taking pleasure in, and being glad that. 'Good' refers to goodness *simpliciter*. Goodness simpliciter should be distinguished from *goodness-for* (as something might be good *for* one's health, or for trees, or for an engine) and from *attributive-goodness* (as in good toasters, good holiday destinations, and good assassins). The buck-passing account is thus that for something (a state of affairs, or perhaps a process or event) to be good simpliciter is for anyone to have sufficient reason to have some pro-response towards that thing. We take the buck-passing account, so understood, to have at least three central attractions.

1.1. Demystifying Goodness Simpliciter

The first attraction of the buck-passing account is that it helps to demystify the notion of goodness simpliciter. Several philosophers, including Peter Geach, Philippa Foot, and Judith Thomson, have argued that no sense can be made of this notion.[6] They insist that talk of goodness is always talk about goodness-for or attributive-goodness. Here is a representative passage from Thomson:

> [A]ll goodness is goodness in a way…if we do not know in what way a man means
> that a thing is good when he says of it 'That's good', then we simply do not know
> what he is saying of it. Perhaps he means that it is good to eat, or that it is good for
> use in making cheesecake, or that it is good for Alfred. If he tells us, "No, no, I meant
> that it is just plain a good thing," then we can at best suppose he is a philosopher
> making a joke.[7]

Geach, Foot, and Thomson each take the claim that there is no such thing as goodness simpliciter to have deep implications for moral philosophy. Most notably, they take it to show that the consequentialist claim that we ought to maximize (final) goodness simpliciter is a non-starter.

The buck-passing account offers a straightforward answer to this worry.[8] On the buck-passing account, being (finally) good simpliciter just is being what there is sufficient reason to value for its own sake. Geach, Foot, and Thomson do not claim to doubt that talk of reasons or of valuing something for its own sake makes sense. And nor should they: it seems plain that some of us do value happiness for its own sake, for example, and it is intelligible to claim that there is sufficient reason to do so.

1.2. Generalising the Account: Specific Values, Good-For, Attributive Good

The second central attraction of the buck-passing account is that it promises to generalise in an attractive way. As we have explained, the buck-passing account is in the first instance an account of goodness simpliciter. However, as we have seen, there are a range of other value properties, such as goodness-for, attributive goodness, and more specific value properties, such as the admirable and amusing. Since it would be surprising if there was no interesting relation between goodness simpliciter and these properties, a promising account of the former should be expected to generalise to these other kinds of value properties, and to illuminate the similarities and differences between them.

The buck-passing account is well-placed to do this. It is easily adapted to offer accounts of specific value properties. The buck-passer can say that something is, for example, admirable if there is sufficient reason to admire it. There are also prospects for generalizing the account to goodness-for and the varieties of attributive goodness. For instance, perhaps Y

is good *for* X if there is sufficient reason for anyone who has reason to care about X to value Y, because they have reason to care about X. And perhaps Y is a good X if there is sufficient reason for anyone who has reason to want an X to value Y, because they have reason to want an X.[9]

These suggestions are rough, but they are enough to indicate that there are serious prospects of developing the buck-passing account so as to generalise to the other value properties and to illuminate the similarities and differences between them.

1.3. Reasons and Value

The third central attraction of the buck-passing account is that it offers straightforward explanations of the various connections between reasons and value. We shall give two examples.

First, the buck-passing account explains what one of us has elsewhere called the *linking principle*.[10] To a first approximation, it seems to be true that:

> If some consideration is a respect in which X is good, then that consideration is a reason to value X.

For example, if another drink would be good insofar as it would bring you pleasure, then the fact that it would bring pleasure is a reason to want another drink; if travel is good insofar as it broadens the mind, then that fact that travel broadens the mind is a reason to want to travel.

The linking principle requires explanation – it is surely not a coincidence that reasons to value and goodness are connected in such a systematic way. The buck-passing account offers a simple explanation of the principle. If for something to be good is for there to be

reasons to value it, then it is plausible that for something to be good *in some respect* is for that respect to *be* a reason to value it. The linking principle thus follows directly from the nature of value.

Second, T.M. Scanlon famously observes that the fact that something is good is not an additional reason to value that thing, over and above the non-evaluative facts which are reasons to value that thing. For example, the fact that a scientific discovery is *good* does not seem to be an additional reason to applaud that discovery, over and above the non-evaluative facts which make the discovery good – for instance, that it sheds light on the causes of cancer.[11] In fact, Scanlon suggests that the goodness of the discovery does not give a reason to applaud it at all. Arguably, this is too strong. However, if the goodness of the discovery is a reason to applaud it, it is not an additional reason in the sense of being a reason which adds any *weight* to the non-evaluative reasons for applauding the discovery.[12]

Again, this observation is straightforwardly explained by the buck-passing account. The goodness of the discovery adds no additional reason to applaud the discovery because the discovery's goodness simply consists in the presence of reasons to applaud it. Thus, it does not add weight to these reasons, even if it is itself a reason.

1.4. The Wrong Kind of Reason Problem

While there is more to say on all these issues, it seems clear that the buck-passing view has some significant prima facie attractions. However, it also faces a very serious problem. Suppose, to give a now well-worn example, that an evil demon threatens to destroy the world unless everyone values him. The demon's threat seems to give everyone sufficient reason to value him – after all, otherwise the world will be destroyed! But the demon's threat does not make him good. The buck-passing account thus seems open to simple counter-examples.[13]

Buck-passers have offered two main sorts of response to this problem. The first response is to reject the apparent counter-examples. Buck-passers who take this line insist that the demon's threat is not a reason to value him. Instead, the threat is merely a reason to want to admire the demon, and perhaps to bring it about that you admire him. On this view, the incentive to admire the demon is like an incentive for having a headache. Such incentives can give you reason to want a headache, and to bring one about, but they do not, strictly speaking, give you reasons for a headache.[14]

The second response is to draw a distinction between two kinds of reasons to value. The *right* kind of reasons to value are considerations which are both reasons to value an object and features that make it good. The *wrong* kind of reasons are considerations which are reasons to value an object but not features that make it good. If a distinction of this sort can be drawn then buck-passers can reformulate their view. They can say that for something to be good is for there to be sufficient *right kind* reason to value it.

Both responses face serious difficulties. The worry about the latter strategy is simple. The most obvious way to draw the distinction between the right and the wrong kind of reasons is in terms of goodness, as above. But, on pain of circularity, the buck-passer cannot draw the distinction in this way. And while several philosophers have now tried to offer non-circular accounts of what makes something a reason of the right kind, it is fair to say that none have been widely accepted. All such accounts seem either to give rise to clear counter-examples or to turn out, on inspection, to rely on the notion of value which the buck-passing account is intended to elucidate. Furthermore, even if the buck-passer can distinguish the right and the wrong kind of reasons in a way that avoids counter-examples, the literature to this point suggests that any such distinction is going to be complicated. This will make the buck-passer's view correspondingly less attractive.[15]

The former strategy is also difficult to defend. As is often pointed out, many find the claim that incentives provide reasons for attitudes clearly true.[16] Proponents of the second strategy thus need to both argue that incentives are not reasons and explain why not. The second of these tasks is particularly challenging.[17] Buck-passers typically accept the reasons-first approach, on which reasons are the basic normative unit.[18] And many accept the further claim that reasons are absolutely basic: that we can give no informative account of what it is for something to be a reason. At most, we can paraphrase: thus Scanlon and Parfit say that reasons are considerations which *count in favour* (or against) a response. These buck-passers have very few resources available to explain why incentives for attitudes are not reasons for those attitudes – since their account of what reasons are is so thin, nothing in it seems to rule this out.[19] For example, suppose that the buck-passer claims that incentives for attitudes are not reasons for attitudes because they fail a *response-condition*, according to which a reason to A must be the sort of thing for which you could A.[20] Even if it could be shown that incentives for attitudes failed this condition, buck-passers are not well placed to endorse the response condition. For, if such a condition holds, either it is a brute truth about reasons, or there is some explanation of *why* it holds. It would be a serious cost for any view to posit the response condition as a brute truth. It is a very strong condition; if true, it cries out for explanation. Suppose, then, that there is an explanation of why this condition holds. The obvious candidate is that it holds because of the nature of reasons. But, if reasons are absolutely basic, it is hard to see how anything about their nature could explain the response condition.[21] Nor does Scanlon and Parfit's paraphrase help. If anything is clear, it is that the fact that the demon will destroy the world unless everyone values him counts in favour of valuing the demon. Counting in favour is not subject to a response condition.

Similar considerations apply to other conditions on reasons that might be invoked here, for example (to anticipate) that reasons must correspond to good reasoning.

We do not want to say that neither of the strategies outlined can succeed, or that the wrong kind of reason problem thus refutes the buck-passing view. Nonetheless, these worries seem serious enough to motivate the consideration of an alternative, non-buck-passing, version of the fitting-attitudes view. In the next section, we outline such a view.

2. Fittingness First

2.1 Fitting-Attitudes Without Buck-Passing

The fitting-attitudes account of value as such does not face the wrong kind of reason problem. Recall that the fitting-attitudes account says that

What it is for X to be good is for it to be fitting to value X.

The demon's threat to destroy the world unless everyone values him might give everyone a reason to value him, but it does not make it fitting to value him – it does not make him fit to be valued. This judgment is widely shared in the literature and we take it to be intuitively clear.[22] Note that it can equally be expressed in other ways. For instance, one might also say that it is not *correct* to value the demon, that one would not be *getting it right* in valuing him, or that the demon is not *worthy* of valuing, or an *appropriate object* of valuing. These expressions can all be used to express the notion of fittingness.

Why isn't it fitting to value the demon? Strictly speaking, all that matters for our purposes at this point is *that* it is not fitting. However, it may be helpful to say something about this, in order to make sure we have satisfactorily individuated the property of fittingness (we will say more about it later). Take a specific valuing attitude, such as

admiration. This attitude sets a *standard* for objects. The standard specifies certain features, having enough (or an appropriate weighted sum) of which makes an object fit to admire— that is, admirable. For example, features like humility and concern for others may be on the list for admiration. The demon is notably lacking in these sorts of features. Thus, it is not fitting to admire him.

The idea of attitudes' having internal standards of fittingness or correctness is a familiar one. The most obvious example is belief. Whatever is the best interpretation of the slogan that belief 'aims at truth', this much is plausible: it is in the nature of the attitude of belief to be assessable as correct or incorrect, according to the truth or falsity of the proposition believed.[23] The attitude of belief thus sets truth as its standard of correctness. We take correctness for belief to be an instance of the broader property of fittingness.[24]

So, the fitting-attitudes account alone does not entail that the demon is good. It does so only when supplemented with the account of fittingness in terms of reasons—the step that turns the fitting-attitudes account into a buck-passing account. Perhaps, then, the defender of the fitting-attitudes account should eschew buck-passing. But we also saw that buck-passing has many attractions. Is the fitting-attitudes account alone equally attractive?

The first such attraction was the demystification of goodness simpliciter. The fitting-attitudes account alone seems sufficient to secure this result. On this approach, goodness simpliciter is understood in terms of what is fitting to value. There doesn't seem to be any particular difficulty in making sense of talk of what is fitting to value.

The second attraction of buck-passing was that it promised to generalise to other value properties, besides goodness simpliciter. The fitting-attitudes account alone does this just as easily. Instead of saying that something is admirable if there is sufficient reason to admire it, we say simply that something is admirable if it is fitting to admire it. Likewise for other specific values. We can distinguish goodness simpliciter, goodness-for, and attributive

goodness in much the same way as the buck-passer does. For X to be good simpliciter, we can say, is for it to be fitting for *anyone* to value X. For Y to be good for X is for it to be fitting for anyone for whom it is fitting to care about X to value Y, because it is fitting for them to care about X. And for Y to be a good X is for it to be fitting for anyone for whom it is fitting to want an X to value Y, because it is fitting for them to want an X. These formulations are rough, but they seem enough to show that the fittingness-first view is in just as good a position as buck-passing when it comes to accounting for these value properties.

The third attraction was the explanation of connections between reasons and value. Here, the fitting-attitudes account alone does not suffice. That's because it has nothing to say about reasons. But perhaps it can explain these connections when supplemented by a plausible account of reasons. We explore this possibility in the next section.

2.2 The Fittingness-First Account of Reasons

How would a suitable account of reasons go? Since we are trying to explain connections between value and reasons, and the fitting-attitudes account understands value in terms of the fittingness of a certain kind of response, it seems sensible to try to understand reasons in terms of fittingness as well.[25] Some ways of doing this are suggested by views that are already in the literature. For example:

For that *p* to be a reason for a response is for that *p* to be evidence that that response is fitting.

Or:

> For that p to be a reason for a response is for that p to help explain why that response is fitting.[26]

However, both of these accounts are subject to forceful objections, parallel to those facing the more familiar claims that reasons are evidence, or explanations, of what you *ought* to do.[27] We propose instead an account of a sort already defended by one of us elsewhere.[28]

As a first step towards this account, note that there seems to be an important connection between reasons and *reasoning*. Reasons, we might say, are *for* reasoning; they are the sort of thing it is suitable to reason with. This suggests that it might be possible to understand reasons as *premises of good reasoning*. Here is a way to make this idea more precise (we will modify this slightly later on):

> For that p to be a reason for a response is for that p to be a premise of a good pattern of reasoning from fitting responses to that response.

To illustrate: that the lights are on is a reason to believe that somebody is home, because there is a good pattern of reasoning with this consideration as a premise, that concludes in the belief that somebody is home; that you promised to meet your friend is a reason to intend to do so, because there is a good pattern of reasoning with this consideration as a premise, that concludes in the intention to meet your friend.

Here, reasoning is understood broadly, as a certain kind of transition in which a set of responses, which we can call premise-responses, leads to some (further) response, which we can call the conclusion-response. This transition is such that the conclusion-response counts as based on, or held in the light of, the premise-responses. The sorts of things that count as responses include, notably, attitudes. For example, theoretical reasoning might involve a

transition from a set of beliefs to a belief. Practical reasoning might involve a transition from a belief and a pro-attitude to some further pro-attitude, to an intention, or, perhaps, to an action.

That p is a premise of a pattern of reasoning when the belief that p, or some other appropriate representation of the consideration that p, is among the premise-responses of that pattern of reasoning.[29]

When we say that reasons are premises of good reasoning, we do not mean that anyone must actually perform this good reasoning. Patterns of reasoning are abstract. They need not be performed in order to be assessable as good or bad reasoning. It may be, however, that in order for the consideration that p to be a reason *for a given subject* to A, that subject must have some of the attitudes that constitute premise-responses of the good pattern of reasoning in virtue of which it is a reason. This would be a way for the view to allow that a subject's attitudes may make a difference to her reasons.[30]

According to our account, reasons are premises of good reasoning *from fitting responses*. The motivation for this restriction is easy to see. Since beliefs are fitting only when true, this restriction ensures that only facts are reasons. It also allows us to avoid further difficulties. For example, suppose that someone were to reason in one of the following ways:

	Belief that the moon orbits the earth
and	Belief that the moon is made of cheese
leads to	Belief that something made of cheese orbits the earth.

	Intention to kill everyone
and	Belief that the only way to kill everyone is to get a nuclear weapon
leads to	Intention to get a nuclear weapon.[31]

These pieces of reasoning instantiate good patterns. It's just that the starting points are inappropriate. Without the restriction to fitting premise-responses, our account would entail that the fact that the moon orbits the earth is a reason to believe that something made of cheese orbits the earth, and it would entail that the consideration that the only way to kill everyone is to get a nuclear weapon is a reason to intend to get a nuclear weapon, at least for someone who intends to kill everyone.

By contrast, consider:

	Belief that the moon orbits the earth
and	<u>Belief that the moon is made of rock</u>
leads to	Belief that something made of rock orbits the earth.

	Intention to prevent the madman from killing everyone
and	<u>Belief that the only way to do so is to disable his nuclear weapon</u>
leads to	Intention to disable the madman's nuclear weapon.

These also look like good pieces of reasoning, but this time the starting points are fitting. Our account therefore entails, plausibly, that the fact that the moon orbits the earth is a reason to believe that something made of rock orbits the earth, and that the fact that the only way to prevent the madman from killing everyone is to disable his nuclear weapon is a reason to intend to disable his nuclear weapon, at least for someone who intends to foil the madman's plan.

At this point, we should say something to clarify the 'good' in 'good reasoning'. This looks like an attributive use of 'good'. Our account of reasons as premises of good reasoning

therefore looks like an account of reasons in terms of attributive goodness – attributive goodness, like other goodness, being understood in turn in terms of fitting attitudes.

We don't see any special problems with construing the view in this way. However, we do not wish to commit to the claim that attributive goodness is prior to reasons. An alternative is to understand the 'good' in 'good reasoning' simply as a placeholder for the coming account of when reasoning is going as it is supposed to. Then, reasons would be understood directly in terms of what it is for reasoning to go in this way. We now turn to providing this account.[32]

Our account begins from reflection on the point of reasoning. Reasoning is a way of revising our attitudes. But the point of reasoning is not merely to change our attitudes – for there are all sorts of ways we can do that. Plausibly, the point of reasoning – what in general makes it worth going in for – is to revise our attitudes so as to get things *right*. That is, the point is to add to our stock of fitting attitudes[33] – for instance, to gain more true beliefs, or intentions for what is worth doing.[34] Of course, reasoning cannot usually generate fitting attitudes from nothing – starting points matter. But if we start from fitting attitudes, we can expect good reasoning to lead us to further fitting attitudes, at least other things equal. This suggests the simple and prima facie plausible view that

For a pattern of reasoning to be good is for it to be fittingness-preserving,

where

For a pattern of reasoning to be fittingness-preserving is for it to be the case that if the premise-responses are fitting, then, other things equal, so is the conclusion-response.

This suggestion constitutes a generalisation of what seems a highly plausible theory of theoretical reasoning, namely the theory that good theoretical reasoning is truth-preserving reasoning. The 'other things equal' clause is required in order to allow for defeasible kinds of reasoning (and so for defeasible reasons). When a good but defeasible pattern of reasoning is defeated, other things aren't equal. For example, consider:

<u>Belief that Jennifer said that it is raining</u>

leads to   Belief that it is raining.

<u>Belief that I promised to meet my friend for lunch</u>

leads to   Intention to meet my friend for lunch.

These ways of reasoning may fail to lead from fitting premise-responses to a fitting-conclusion response, for example because Jennifer is lying or because I can save someone's life by missing lunch. But they nonetheless exemplify good pattern of reasoning. That's because a situation in which testimony is unreliable, or in which I can save someone's life by breaking a promise, is one in which other things are not equal.[35]

Though initially plausible, this suggestion faces problems. Consider the following examples:

<u>Belief that grass is green</u>

leads to   Belief that 79 is prime.

<u>Belief that reliable Jane told me that Jill is admirable</u>

leads to   Admiration of Jill.[36]

The reasoning in these examples is fittingness-preserving but it is not clear that it is good reasoning. And it is even less clear that the fact represented by the premise-belief, in these cases, constitutes a reason for the conclusion-response. So, having the property of being fittingness-preserving may not be sufficient for a pattern of reasoning to be good, and being a premise of a fittingness-preserving pattern of reasoning seems insufficient for being a reason.

We will therefore refine our account of good reasoning. In order to do so, we start from the observation that, when you perform a piece of reasoning that preserves fittingness, you may or may not be *sensitive* to its having this property. This is a familiar point. There is a difference between, for example, inferring a conclusion *because* it follows from your premises, and inferring the same conclusion from the same premises because you mistakenly think that *anything* follows from those premises. In the latter case, although the transition you make happens to be okay, something is still going wrong with your reasoning. Something is going wrong because you lack sensitivity to the appropriateness of the transition.

We can say that you are not reasoning *competently*, in the latter case. Here, we use the term 'competence' in a stipulative sense: you reason competently just in case your reasoning instantiates a fittingness-preserving pattern, and you perform it in a way that is sensitive to this.

Now, consider the first problematic example above. The problem with this reasoning is that the fact that grass is green is irrelevant to whether 79 is prime: although the transition is fittingness-preserving, this is just because it is necessarily true that 79 is prime. The premise could be anything at all. But this means that, in order to be sensitive to the fittingness-preserving character of the transition, you must already be aware that 79 is necessarily prime. So, you can't come to believe that 79 is prime by competently performing this reasoning. In order to perform it competently, you would have to already believe this.

This, we suggest, is what prevents the pattern from counting as a good one. We therefore propose that

> For a pattern of reasoning to be good is for it to be the case that one can come to hold the conclusion-response by competent reasoning from the premise-responses.

Here we are simply adding a condition to the account proposed earlier: good patterns are not only fittingness-preserving, but also such that one can be sensitive to this property in using them to add to one's stock of attitudes. This is in the spirit of the earlier proposal. Patterns that can't guide us in the acquisition of fitting attitudes are *useless*.

This account relies on the notion of being *sensitive* to fittingness-preservation. The need for a distinction between making a transition that happens to be okay, and doing so in a way that is sensitive to its being okay, should be uncontroversial. Nonetheless, one might wonder what such sensitivity involves. The most straightforward way to understand sensitivity would be in *causal* terms: for your reasoning to be sensitive to fittingness-preservation is for it to be appropriately caused by this property.[37] Alternatively, it could be understood in terms of how you would reason in certain counterfactual circumstances, for example circumstances involving differences in the premise-responses. While we lack the space to develop these suggestions here, it is clear that neither of them need introduce circularity into our account by appealing to something normative other than fittingness.

What about the transition from believing that reliable Jane told me that Jill is admirable to admiring Jill? To deal with this, we want to make a further distinction: between *basic* and *complex* patterns of reasoning. Once again, this is a distinction that any plausible view of reasoning must allow for. There is a difference between a direct transition from some set of attitudes to a further attitude—for example, a *modus ponens* transition—and an indirect

transition, in which several steps occur—for example, a transition consisting of several *modus ponens* steps. We can capture this distinction as follows:

> For a pattern of reasoning to be a good *basic* pattern is for it to be the case that one can come to hold the conclusion-response directly by competent reasoning from the premise-responses.

To say that one can come to hold the conclusion-response *directly* is to say that there are no intermediary responses, such that they must be held on the basis of the premise-responses, and the conclusion-response held on the basis of the intermediary ones, in order for one to be reasoning competently.

The transition from believing that reliable Jane told me that Jill is admirable to admiring Jill does not instantiate a good basic pattern. To be good reasoning, it must, implicitly at least, involve two steps:

> Belief that reliable Jane told me that Jill is admirable

leads to   Belief that Jill is admirable.

> Belief that Jill is admirable

leads to   Admiration of Jill.

After all, if I am not at least implicitly 'passing through' the belief that Jill is admirable—that is, if I don't actually believe what Jane told me—then I am not sensitive to the fittingness-preserving character of the transition from believing that Jane told me this to admiring Jill.

Now it is time for the slight modification, promised earlier, of our account of reasons:

For that *p* to be a reason for a response is for that *p* to be a premise of a good *basic*

pattern of reasoning from fitting responses to that response.

With this modification, we get the result that the fact that reliable Jane said that Jill is

admirable is not itself a reason to admire Jill. The fact that Jill is admirable *is* such a reason.

But this seems plausible, since it does not entail that this fact is a reason to admire her over

and above the non-evaluative facts that make her admirable (we return to this issue below).

Is this modification *ad hoc*? We don't think so. Given the distinction, which any

plausible view of reasoning should allow, between basic and complex patterns of reasoning,

our view faces the question of whether reasons correspond to *all* good patterns of reasoning,

or just the basic ones. It seems legitimate for reflection on examples – such as the case of

reliable Jane – to guide our answer to this question. Furthermore, to the extent that more

theoretical pressures come to bear here, they seem to favour our restriction. Reasons for a

response *directly* support that response, and they should be able to guide us through our

sensitivity to that support. It is thus natural to think that reasons for a response must be such

that you can competently reason from them *directly* to that response. It would be an odd sort

of reason that, in reasoning directly from it to the response it supports, you were *guaranteed*

to be reasoning incompetently. If a pattern is non-basic, however, it is not possible to

competently reason directly from the premise-responses to the conclusion-response. The

guidance role of reasons thus provides a principled motivation for the restriction to basic

patterns.[38]

Let us summarise the view we are proposing. What is good is what is fitting to value.

Reasons are also explained in terms of fittingness. Reasons are premises of good patterns of

reasoning. That is, they are premises of fittingness-preserving patterns to which one can be sensitive. Thus, the basic normative unit is fittingness. We put *fittingness first*, and explain other normative properties in terms of fittingness together with certain other properties, such as those of the attitudes of valuing.

Of course, there is a great deal more to be said about the details of our fittingness-first view. In particular, the 'other things equal' clause and the notion of sensitivity to fittingness-preservation call for further elaboration. But we hope to have done enough to suggest that there's a promising view in the offing, and a legitimate rival to buck-passing.

2.3 Fittingness First v. Buck-Passing

Now that we have outlined an account of reasons in terms of fittingness, we are in a position to show how the fittingness-first view captures the third central attraction of buck-passing – its explanation of connections between reasons and value.

The first connection between reasons and value that we considered was the linking principle, according to which respects in which X is good are reasons to value X. Our account seems to entail the linking principle. The transition from believing that X is F, where F is some respect in which X is good, to valuing X, constitutes good reasoning. For example, the transition from believing that the sculpture is elegant to admiring the sculpture preserves fittingness, and can be performed in a way that is sensitive to this fact. Thus the fact that the sculpture is elegant is a reason to admire it.

What about Scanlon's observation that the fact that something is good is not an additional reason to value that thing, over and above the non-evaluative facts which are reasons to value that thing? Our account allows that something's being good is *a* reason to value it. But it also accommodates the observation that this isn't a *further* reason to value the

thing – that is, a reason that adds any *weight* to the reasons provided by the non-evaluative facts. We take it that a reason R adds weight to a set of reasons S just when there is a possible opposing set of reasons O that is outweighed by R and S together, but not by S alone. On our view, this means it is good reasoning to move from R and S, together with O, to a response, whereas it is not good reasoning to move from S and O to that response.[39] The goodness of something cannot in this way add weight to the non-evaluative reasons to value it. That's because the total non-evaluative facts are sufficient for goodness – for it to be fitting to value the thing.[40] So, the transition from belief in these facts to valuing the thing will be fittingness-preserving. Adding the belief that the thing is good will be redundant – it will not yield a good pattern of reasoning that was not already available.

Thus, the fittingness-first view has all of the attractions of the buck-passing account but avoids the wrong kind of reason problem. And that's not all. The fittingness-first view also offers some further attractions. We saw that buck-passers, in response to the wrong kind of reasons problem, sometimes try to draw a distinction between right-kind and wrong-kind reasons. Our account delivers a principled distinction of this sort. In particular, it entails that wrong-kind reasons aren't reasons, and does so in a plausible way. The transition from believing that the demon threatened to destroy the world unless everyone values him to valuing the demon does not preserve fittingness. So the threat is not a reason to value the demon. However, it will be good reasoning to move from believing that the demon made this threat to *wanting* to value him, intending to bring it about that one values him, and so on. So the threat is, plausibly, a reason for these latter attitudes. Philosophers who have judged wrong-kind reasons to be reasons for the target attitudes have thus been misled by the fact that they are reasons for closely related attitudes.

This treatment of wrong-kind reasons is plausible, but it is not available to the buck-passer. This is not to say that the buck-passer can't claim that wrong-kind reasons aren't

reasons. The point is, rather, that their account lacks the resources to justify and explain this claim. Insofar as buck-passers either give no account of what reasons are, or say merely that reasons are considerations which count in favour, there is nothing in their account to support the claim that putative wrong-kind reasons aren't reasons. By contrast, an attractive explanation of why they are not falls directly out of our account.

Finally, our account explains more than the buck-passing account. It not only explains the connections between value and reasons, but also the connections between reasons and reasoning,[41] and the connections between all of these things and fittingness. It explains why you can't have reasons for things like headaches, which cannot be fitting and cannot be conclusions of reasoning. Relatedly, it explains the plausible *response condition* on reasons, mentioned earlier: if a consideration is such that you can't respond to it by A-ing, then *a fortiori* you can't competently reason from it to A-ing.[42] So, on our account, such a consideration is not a reason to A. In addition, our account says what all good reasoning has in common, in a way that respects the plausible idea that truth-preserving deduction is a paradigm of it.

3. Objections and Replies

We have now outlined our fittingness-first view and argued that it shares all the attractions, and more, of the buck-passing view, while avoiding the central objection to buck-passing. In this section, we consider several objections to our view.

3.1. What is Fittingness?

We are not the first to note that proponents of the fitting-attitudes account of value can avoid the wrong kind of reason problem by rejecting the buck-passer's account of fittingness in terms of reasons. As Rabinowicz and Rønnow-Rasmussen observe in their seminal discussion of the wrong kind of reason problem, Brentano and Ewing – the founding fathers of contemporary fitting-attitudes accounts – took fittingness as basic and would reject the claim that the evil demon's threat makes it fitting to value him.[43] Nonetheless, Rabinowicz and Rønnow-Rasmussen object to this proposal:

> [H]ow are we supposed to clarify this basic relation of fittingness? If we leave it as primitive, then FA analyses can be accused of reducing the concept of value to another notion that is at least as troublesome. On the other hand, if we try to explicate this relation, we seem to be drawn to something like this: a pro- or contra-attitude is fitting with regard to its object if and only if it is adequate to the value of that object. Thus, the analysis moves in a circle: from value to fittingness and from fittingness to value.[44]

Our view avoids the second horn of this dilemma. We take fittingness as basic, rather than analysing it in terms of value. So our view is not circular. Presumably then, Rabinowicz and Rønnow-Rasmussen's objection would be that we reduce value to something 'at least as troublesome'. Now it is not entirely clear what they have in mind by this. But part of the worry seems to be that the notion of fittingness is obscure. If so, then, pending further clarification, it cannot help to shed light on value or, as we have suggested, on reasons. Thus understood, the objection takes the form of a challenge: can we say more about what fittingness is?

This challenge has to be handled with some care. If the demand is for a reductive account of fittingness, then we're not going to provide one, since we take fittingness to be normatively basic, and have no project to reduce the normative to the natural or descriptive (we are neutral on whether such a reduction is possible). However, a reductive account is not the only way to identify or characterise some phenomenon.

We can begin by reiterating that fittingness does seem to be a distinctive normative notion. The wrong kind of reason problem itself is evidence of this. It seems to be a common judgment that the demon's threat provides reasons to value him but does not make it fitting to value him. Judgments about what you ought to do and what is permissible to do seem to go with reasons in these cases. If you think that the demon's threat provides everyone with reason to value him, then surely you will also think that everyone ought, and so is permitted, to admire him. The wrong kind of reason problem shows that we can and do distinguish the notion of fittingness from other related notions.

It might be objected that this is so only because we think of 'fittingness' in terms of value – we think of 'fitting' as meaning 'fitting to the value of the object' (this seems to be what Rabinowicz and Rønnow-Rasmussen have in mind above). But other versions of the wrong kind of reason problem show that this is not right. If an evil demon will kill you unless you believe that the moon is made of cheese, then it is tempting to think that this gives you reason to believe that the moon is made of cheese, but it does not make it fitting to believe this. But this cannot be because talk of fittingness is talk of what's fitting to the value of an object. The object of a fitting belief need not be valuable.[45]

But even if it is granted that there is a distinctive ordinary notion of fittingness, why think that fittingness is normatively basic? Well, in part we think that this supposition should be judged by its fruits. To the extent that the picture of the normative and evaluative domain sketched in this paper is attractive, that is evidence that fittingness is first. But we can also

provide some further motivation by briefly considering three proposals for how fittingness might be reduced to something else.

The first proposal is that fittingness should be analysed in terms of reasons: what is fitting is what there is sufficient reason for. However, this proposal would face the same problem as the view that goodness should be understood in terms of reasons: the wrong kind of reason problem. This paper is motivated in large part by the challenges buck-passers face trying to solve this problem. We won't reconsider these challenges here.

A second proposal is that fittingness is a matter of correct or accurate representation: an attitude is fitting if it correctly represents its object.[46] This proposal works well for belief – beliefs are fitting when they correctly represent how things are. The problem is that the proposal accommodates the full range of attitudes that can be fitting only by making highly contentious assumptions about the nature of attitudes.[47] It is clear that attitudes such as admiration, regret, hope, desire, and intention can be fitting. But it is not at all clear that these attitudes represent their objects in such a way that they are fitting just when their objects are as they are represented to be. Perhaps, for instance, envy represents another as having something desirable which you lack, desire represents its object as good, and intention represents an action as worthy of pursuit. If these assumptions hold then the proposal would have plausible implications about when these attitudes are fitting. But assumptions of this sort are a very big commitment for an account of fittingness to take on. It would be surprising if the possibility of assessing the fittingness of these attitudes turned on them. Furthermore, these assumptions would appear to rule out a fitting-attitudes account of value, by rendering it circular.[48]

A third proposal is that fittingness is a matter of permissibility: for it to be fitting to A is for it to be permissible to A.[49] This proposal also has something going for it. Fittingness seems more like a permissive notion than a requiring one – you're not required to believe

every truth, or value everything that's valuable – and, like permissibility, it is an overall notion rather than a pro tanto one. Nonetheless, the view does not seem satisfactory. We have already seen some evidence for this. In wrong kind of reason scenarios, those who find it intuitive that there is reason to value the demon will also find it intuitive that it is permissible to value the demon. This suggests that we can distinguish our judgments about fittingness from our judgments about permissibility.[50]

There are also further grounds for taking there to be a distinction here. A distinction between fittingness and permissibility follows from what we have already argued, given the familiar idea that permissibility should be understood in terms of reasons:

> For it to be permissible for you to A is for you to have at least as weighty reason to A as to not A.

We have argued that fittingness cannot be understood in terms of reasons. Thus, fittingness and permissibility are distinct.[51]

We take the above to go a fair way to meeting Rabinowicz and Rønnow-Rasmussen's challenge. There are good grounds for thinking that fittingness is distinct from other normative and evaluative notions and that it cannot easily be analysed in other terms.

So far in this section we have talked about what fittingness is not. Is there anything positive to be said about what it *is*? As we said in section 2.1 above, the fittingness of a response is a matter of the satisfaction a certain standard, internal to that response. One way of getting more of a grip on the sort of property fittingness is might be to say something about how a type of response gets to have a standard of fittingness. Here we offer a very speculative suggestion about how this might work. This suggestion is strictly an optional add-on; nothing in the main argument of the paper, or in the fittingness-first view as such, depends on it.[52]

Let's take the case of belief first. The belief that *p* is assessable as fitting or not depending on whether *p* is true or false. But what makes this the case? We suggest that this has to do with the constitutive functional role of belief, and in particular its role in reasoning. The belief that there is beer in the fridge, for example, might lead through some simple reasoning to the desire to go to the fridge, to the intention to get a glass, and so on. Crucially, the fittingness of these further responses will normally depend on whether the initial belief is true. Whether it is fitting to want to go to the fridge, or whether it is fitting to intend to get a glass, depends on whether there really is beer in the fridge.

So, belief by its nature tends to lead to certain further responses whose fittingness systematically depends on the content of the belief having a certain feature—namely, truth. We suggest that this is why belief has a standard of fittingness, and this is why the standard is truth.

A similar story, we suggest, can be told for other attitudes that have standards of fittingness. Consider desire. This attitude has a role in reasoning to further attitudes: the desire for some beer might lead you to intend to go to the fridge, to regret not finding beer there, and so on. Whether it is fitting to intend to go to the fridge, or to regret not finding beer there, depends on certain features of the prospect of drinking beer—whether it would be pleasant, whether it would be unhealthy, and so on. And these are just the sorts of features that, plausibly, contribute to the standard of fittingness for desire: something's being pleasant tends to make it fitting to desire, whereas its being unhealthy tends to make it unfitting to desire. So, the features that make desire fitting are those on which depends the fittingness of the further responses that desire leads to.

While it is clear enough what feature makes a proposition fitting to believe, it is far from obvious which non-evaluative features make something desirable, or in what measure. But what we say above indicates how a given non-evaluative feature, such as that of being

pleasant, might get to be among those that make something desirable. Perhaps it would be possible to give an exhaustive account of these features, and of the contribution each makes to desirability, but such an account might be highly complex, and providing one would require substantive first-order normative theorising. The same is likely to be true for many other responses.

Generally, we suggest that something like the following is true. A response can get to have a standard of fittingness when that response, by its nature, leads through reasoning to further responses whose fittingness depends on the response's content or object having certain features. This is not a reductive account of fittingness. What's more, it takes for granted facts about what makes certain responses fitting in accounting for what makes other responses fitting.[53] Nonetheless, we suggest it can still help us to understand the kind of property fittingness is.

Note that this proposal helps to explain why it's not fitting to admire the demon. Admiration of someone leads to further attitudes like taking pleasure in their achievements and desiring to emulate them. These attitudes would not be fitting when directed towards the demon.

Naturally, much more would need to be said to clarify and defend this speculative proposal. We offer it here merely as an optional suggestion for how fittingness might be elucidated in a non-reductive way.[54]

3.2. The Wrong Kind of Fittingness?

Here is a second objection to our account:

Even if fittingness is distinct from reasons, oughts, permissibility, and the like, the fittingness view has no real advantage over the buck-passing view. Just as the buck-passing view faces the wrong kind of reason problem, the fittingness view faces the 'wrong kind of fittingness' problem. For example, when the demon threatens to destroy the world unless you admire him, it would be *prudentially fitting* to admire the demon, even if it would not be fitting in the sense of satisfying admiration's internal standard. The fittingness view must thus distinguish between these different fittingness properties. But this task looks no easier than the buck-passer's wrong kind of reason problem.

In order to assess this objection, it is helpful to recall how the WKR problem arises for the buck-passer. It starts with cases where there seem to be reasons for a pro-attitude, like admiring the demon, without the object of the attitude being thereby valuable. These cases appear to be counterexamples to the buck-passing account. Their force depends on the intuition that the demon's threat (for example) really does give a reason to admire him. As we noted, many find this intuition compelling. And it can be bolstered by further considerations. It is often claimed that a reason for a response is just something that counts in favour of it; this approach, as we noted, is particularly natural for the buck-passer. The demon's threat clearly counts in favour of admiring him. Furthermore, the ordinary notion of a reason includes the notion of a prudential reason. In so far as the demon's threat makes it prudent to admire him, it seems to give a prudential reason to do so.

To deal with these apparent counterexamples, the buck-passer must either reject the intuition that the demon's threat gives a reason to admire him, or refine her account by distinguishing between the right and wrong kind of reasons. Either way, as we explained, she acquires burdens.

To say that there is a 'wrong kind of fittingness' problem is to say that a corresponding line of reasoning applies with respect to fittingness. But in fact the corresponding line of reasoning fails at the first step. While it is intuitive that the demon's threat gives a reason to admire him, it is not intuitive that the threat makes admiration of the demon fitting. As we noted earlier, this judgment is widely shared in the literature. And it coheres with intuitive judgments about similar cases: there's little temptation to think that the need to ingratiate yourself with your boss makes her jokes worthy of amusement, or that Pascal's wager makes belief in a god correct. Moreover, further considerations back it up. The notion of a response's being fitting is not the same as the notion of something's counting in favour of it. Nor does it have the notion of a prudential kind of fittingness built into it. So, in denying that it is fitting to admire the demon, we do not acquire a burden to defend and explain this verdict, in the way that the buck-passer who denies that there is a reason to admire the demon does.

Of course, a full defence of our view should include some elucidation of fittingness that helps to vindicate and explain intuitive verdicts, such as that admiration of the demon is not fitting. We have offered some of this in sections 2.1 and 3.1.

Perhaps the buck-passer's options aren't exhausted. As we explained, one move she can make is to distinguish between the right and wrong kind of reasons. The problem is then to explain this distinction without presupposing value. But the buck-passer might claim that she can put her view and ours back on all fours by taking the property of being a *reason of the right kind* as basic. Reasons of the wrong kind may then be understood in terms of right-kind reasons – for instance, perhaps wrong kind reasons are just right-kind reasons to favour an attitude.[55] The distinction between the right and wrong kind of reasons does after all have some intuitive purchase, and it's clear enough that the demon's threat falls on the wrong side. This buck-passer thus avoids a counterintuitive verdict about the demon case, and, while she

does appeal to a normative notion that is taken as basic, this is no disadvantage compared to our view, which does the same.

The problem with this suggestion is that it's hard to see how the property of being a reason of the right kind could be normatively basic. The obvious way to understand this property is conjunctively: being a right-kind reason is being a reason, plus meeting some further condition. But so understood, the property of being a reason must be prior to that of being a reason of the right kind.[56] By contrast, fittingness does not seem to be a matter of having some normative property plus meeting some further condition. It is thus a more plausible candidate for being normatively basic.

Where does this leave 'prudential fittingness'? This seems to be just another term for 'prudence'. We do not deny that prudence is a standard of evaluation that can be applied to attitudes and other responses. But not all standards are standards of fittingness. We have no difficulty in distinguishing between fittingness and other standards of evaluation, as the widely shared judgments noted above illustrate.[57] Of course, for any arbitrary standard of evaluation, we could call things that meet that standard 'fitting' with respect to it, but this wouldn't create a philosophical problem of distinguishing this kind of 'fittingness' from the property we have been talking about here - a distinct, normative property, familiar in ordinary thought.

3.3. Pro Tanto Goodness

Here is a third objection:

> According to the fitting-attitudes account, something is good if it is fitting to value it. But this is not always so. An outcome may be good in some respect even if, because it

is very bad in other ways, it is not fitting to value that outcome. For example, consider an outcome in which cats live happy lives while all humans spend their lives in agony. This outcome is good in one respect – the cats are happy. But it is not fitting to value it – for instance, to wish that this outcome obtained, or to hope to bring it about – because there is so much suffering in it.[58]

This objection may seem to favour the buck-passing account over the fitting-attitudes account. On the buck-passing view, we can say that this outcome is good in one respect, because the fact that the cats are happy is a reason to value this outcome. Since this does not entail that it is fitting to value the outcome, this does not seem implausible. More generally, it seems an attraction of buck-passing that it understands pro tanto goodness (goodness *in some respect*) in terms of pro tanto reasons. Conversely, the problem for the fitting-attitudes account seems to be that, since it is fitting to value an outcome only if it is good overall, the fitting-attitudes account cannot give an account of pro tanto goodness.

We believe that this objection can be met. To do so, we need to distinguish between pro tanto and overall pro-attitudes. A pro tanto pro-attitude is an attitude towards something in some respect. An overall pro-attitude is an attitude towards something overall. These things can come apart. Thus I might admire someone in respect of their courage and ingenuity, while disapproving of the use they put these traits to and so not admiring them overall.

Given this distinction, it is natural to understand pro tanto goodness in terms of the fittingness of pro tanto pro-attitudes. We can say that for X to be good in respect R is for it to be fitting to value X in respect R. In the example above, it is not fitting to value the outcome in which everyone suffers but some cats are happy overall. But it is fitting to value it in respect of the cats' happiness. For instance, if you knew of such an outcome it would be

fitting to be glad of it insofar as the cats are happy, while being sorry overall. This implication of the account does not seem implausible.

This account of pro tanto value also helps us to answer an objection to fittingness-first views due to Mark Schroeder. Schroeder claims that there are grounds for thinking that fittingness is to be understood in terms of reasons which parallel the grounds for thinking that 'ought' is to be understood in terms of reasons. We should think 'ought' is to be understood in terms of reasons because this explains why what you ought to do is affected by the way in which competing reasons weigh up. For instance, you ought to keep your promise to meet your friend if this means missing your favourite TV show but not if it means leaving someone to die. Examples like this suggest that what you ought to do is determined by the balance of reasons. Schroeder suggests that, in a similar way, what is fitting is determined by the balance of competing considerations:

> If Mary is more steadfast than Claire, then other things being equal – if that is the only relevant consideration in play – she is also more admirable – that is, it is fitting to admire her more. But if Claire is far more accomplished than Mary, then it may be that Claire is more admirable than Mary – that is, it may be that it is fitting to admire Claire more. The fact that Mary is more steadfast than Claire is a reason to admire her more, and the fact that Claire is far more accomplished is a reason to admire her more. Which one it is fitting to admire more is affected by each of these reasons, and by their relative weights. Finally, we should understand claims of fittingness to be claims about the balance of (the right kind of) reasons, precisely because that best explains why their truth is affected in the way that it is by the balance of (the right kind of) reasons.[59]

However, given what we have said about pro tanto value properties, we can offer a different explanation of the way in which claims of fittingness are determined by competing considerations. On our view, the fittingness of valuing something overall will be determined by the fittingness of valuing it in certain respects. In Schroeder's example, it is fitting to admire Mary in respect of her steadfastness and fitting to admire Claire in respect of her accomplishments. Furthermore, valuing attitudes come in degrees – it will be fitting to admire Claire to a certain degree in respect of her accomplishments and fitting to admire Mary to a certain degree in respect of her steadfastness. This observation allows us to give a simple account of how the fittingness of admiring one more than the other is determined by the fittingness of admiring each in certain respects. It will be fitting to admire Claire overall more than Mary if it is fitting to admire Claire to a greater degree in respect of her accomplishments than it is to admire Mary in respect of her steadfastness. While this explanation of how fittingness can be determined by competing considerations is somewhat schematic, it seems no more so than the alternative Schroeder offers.[60]

3.4. Overkill?

Here is a fourth objection:

> The fittingness-first view is overkill. The view of reasons that you defend – that reasons are premises of good reasoning – implies that putative wrong kind of reasons are not reasons. Buck-passers can thus defend their view just by accepting this account of reasons. There is no need to give up the buck-passing account or accept that fittingness is basic.

The problem with this objection is that the view it outlines looks both circular and implausible. On the view outlined, value and fittingness are understood in terms of reasons, and reasons are understood in terms of good reasoning. Such a view looks circular because the view that reasons are premises of good reasoning makes reference to fittingness:

> For that p to be a reason for a response is for that p to be a premise of a good basic pattern of reasoning from fitting responses to that response.

As we saw, this reference to fittingness is necessary – without it the view has implausible consequences. But if fittingness is understood in terms of reasons, rather than taken as basic, then this account of reasons is circular.

The view looks implausible because it takes the basic normative unit to be that of a good pattern of reasoning. But good patterns of reasoning do not look like good candidates to be the basic normative unit. It is plausible that the point of reasoning is to take us somewhere worth going – for instance, to expand what we know, come to intend what is worth doing, or, as we would say, to lead us to fitting responses. And if this is the point of reasoning, it seems very plausible that good patterns of reasoning should be understood as patterns which serve this aim. If good reasoning is in this way instrumental, it is hard to see how good patterns of reasoning could ground of all normativity.

We thus take it that there are ample grounds for preferring the fittingness first view to the version of the buck-passing account this objection suggests.

4. Conclusion

Buck-passing, despite its attractions, faces the wrong kind of reason problem. The fittingness-first view does not face this problem. As we saw, the problem arises when the fitting-attitudes account of value is supplemented with an account of fittingness in terms of reasons. The fittingness-first view does not take this step. But as we have also seen, the fittingness-first view also has all the attractions, and more, of the buck-passing account. We thus take there to be a strong case for proponents of fitting-attitudes accounts to accept the fittingness-first view, rather than the buck-passing account.

Of course, much remains to be done for a full vindication of the fittingness-first view. While we have presented several ways in which we take the fitting-attitudes account, and the other elements of our view, to be attractive, we have not developed these suggestions into full arguments for the view, and we have left many promissory notes along the way. For example, a full defence of a fitting-attitudes account would need to do much more to show that the account of goodness simpliciter can be satisfactorily developed into an account of goodness-for and attributive goodness. Our account of reasons is also little more than an outline – a full defence would require us to say much more about the crucial 'other things equal' clause in the account of good reasoning, and about what it is to be sensitive to fittingness-preservation. Finally, of course, there is more that might be said about the central notion of fittingness; in particular our tentative suggestion that attitudes might come to have standards of fittingness in virtue of their functional role requires further development. These omissions notwithstanding, we hope to have done enough to show the attractions of taking fittingness first, and thus to have made these remaining questions look worth pursuing.

---

[1] Franz Brentano, *The Origin of Our Knowledge of Right and Wrong* (Oxford: Routledge, 1889/2009); A.C. Ewing, *The Definition of Good* (London: Routledge & Kegan Paul, 1948). It is also suggested by, e.g.,

Henry Sidgwick, *The Methods of Ethics, 4th edition* (London: Macmillan, 1890); C.D. Broad, *Five Types of Ethical Theory* (London: Routledge & Kegan Paul, 1930); Richard Brandt, "Moral Valuation," *Ethics* 56 (1946): 106–21; John McDowell, "Values and Secondary Qualities," reprinted in his *Mind, Value and Reality* (Cambridge, Mass: Harvard University Press, 1998); Roderick Chisholm, *Brentano and Intrinsic Value* (Cambridge: Cambridge University Press, 1986); David Wiggins, "A Sensible Subjectivism?" in his *Needs, Values, Truth: Essays in the Philosophy of Value* (Oxford: Blackwell, 1987). For more on the history of fitting-attitudes accounts, see Wlodek Rabinowicz and Toni Rønnow-Rasmussen, "The Strike of the Demon: On Fitting Pro-Attitudes and Value," *Ethics* 114 (2004): 391-423.

[2] Cf. Derek Parfit, *On What Matters* (Oxford: Oxford University Press, 2011); T.M. Scanlon, *What We Owe To Each Other* (Cambridge, MA: Harvard University Press, 1998); Mark Schroeder, *Slaves of the Passions* (Oxford: Oxford University Press, 2007); John Skorupski, *The Domain of Reasons* (Oxford: Oxford University Press, 2010).

[3] For discussion of how to understand metaphysical accounts of properties, see, e.g. Kit Fine, "Essence and Modality," *Philosophical Perspectives* 8 (1994): 1-16; Frank Jackson, *From Metaphysics to Ethics* (Oxford: Oxford University Press, 1997); Schroeder, *Slaves of the Passions*, 61-72; Ralph Wedgwood, *The Nature of Normativity* (Oxford: Oxford University Press, 2007), 136-44.

[4] For alternative, 'no priority', views about the structure of the normative and evaluative domain, see McDowell, "Values and Secondary Qualities", Wiggins, "A Sensible Subjectivism?".

[5] Buck-passers will differ on the precise formulation. These details won't matter here.

[6] Peter Geach, "Good and Evil," *Analysis* 17 (1956): 32-42; Philippa Foot, "Utilitarianism and the Virtues," *Mind* 94 (1985): 196-209; Judith Jarvis Thomson, "The Right and the Good," *Journal of Philosophy* 94 (1997): 273-98, *Normativity* (Chicago and La Salle: Open Court, 2008).

[7] "The Right and the Good", 276.

[8] Cf. Michael Zimmerman, *The Nature of Intrinsic Value* (Lanham, MD: Rowman and Littlefield, 2001).

[9] Cf. Richard Rowland, "Reasons as the Unity among the Varieties of Goodness," *Pacific Philosophical Quarterly* (forthcoming); Mark Schroeder, "Value and the Right Kind of Reason," in *Oxford Studies in Metaethics, Volume 5*, ed. Russ Shafer Landau (Oxford: Oxford University Press, 2010), 25-55; Skorupski, *The Domain of Reasons*.

[10] Jonathan Way, "Value and Reasons to Favour," in *Oxford Studies in Metaethics, Volume 8*, ed. Russ Shafer-Landau (Oxford: Oxford University Press, 2013) 27-49.

[11] *What We Owe*, 97.

[12] Cf. Mark Schroeder, "Buck-Passers' Negative Thesis," *Philosophical Explorations* 12 (2009): 341-7; Parfit, *On What Matters*.

[13] For further examples, see, e.g. Rabinowicz and Rønnow-Rasmussen, "The Strike of the Demon"; Justin D'Arms and Daniel Jacobson, "Sentiment and Value," *Ethics* 110 (2000): 722-48; "The Moralistic Fallacy: On the "Appropriateness" of Emotions," *Philosophy and Phenomenological Research* 61 (2000): 65-90; Schroeder, "Value and the Right Kind of Reason".

[14] Skorupski, *The Domain of Reasons*; Jonathan Way, "Transmission and the Wrong Kind of Reason," *Ethics* 122 (2012):489-515.

[15] The best known account of the distinction between the right and the wrong kind of reasons is the 'object-given/state-given' theory of Parfit, *On What Matters*; Christian Piller, "Content-Related and Attitude-Related Reasons for Preferences," *Philosophy* 59 (2006): 151-81; Phillip Stratton-Lake, "How to Deal with Evil Demons: Comment on Rabinowicz and Rønnow-Rassmussen," *Ethics* 115 (2005): 788-98. Objections to this view are offered by Daniel Jacobson, "Fitting Attitude Theories of Value," *Stanford Encyclopedia of Philosophy*, ed. E.N. Zalta, http://plato.stanford.edu/entries/fitting-attitude-theories/ (2011); Rabinowicz and Rønnow-Rassmussen, "The Strike of the Demon"; Mark Schroeder, "Value and the Right Kind of Reason", "The Ubiquity of State-Given Reasons," *Ethics* 122 (2012): 457-88. Other accounts include those of Pamela Hieronymi, "The Wrong Kind of Reason," *Journal of Philosophy* 102 (2005): 437-57; D'Arms and Jacobson, "The Moralistic Fallacy"; Sven Danielsson and Jonas Olson, "Brentano and the Buck-Passers," *Mind* 116 (2007): 511–522; Schroeder, "Value and the Right Kind of Reason"; Nathaniel Sharadin, "Reasons Wrong and Right," *Pacific Philosophical Quarterly* (forthcoming). Worries about D'Arms and Jacobson's and Schroeder's views are noted below (section 3.1). Hieronymi's account is not intended to help the buck-passer (453-56). Danielsson and Olson's and Sharadin's accounts rely on a notion of fittingness (they use the term 'correctness') – their views thus seem to be in tension with the reasons-first approach and closer to that defended here.

[16] D'Arms and Jacobson, "Sentiment and Value", "The Moralistic Fallacy"; Rabinowicz and Rønnow-Rassmussen, "The Strike of the Demon"; Danielsson and Olson, "Brentano and the Buck-Passers".

[17] E.g. Way, "Transmission and the Wrong Kind of Reason", provides an argument that wrong-kind reasons are not reasons, but does not attempt an explanation of why they are not.

[18] They need not: buck-passers could instead explain reasons in terms of some other normative property. If they appeal to fittingness then their view may not be so different from ours. We will consider another version of this strategy in 3.4.

[19] Hieronymi, "The Wrong Kind of Reason," 442.

[20] Cf. Bernard Williams, "Internal and External Reasons," reprinted in his *Moral Luck* (Cambridge: Cambridge University Press, 1981); Nishi Shah, "A New Argument for Evidentialism," *The Philosophical Quarterly* 56 (2006): 481–498; Parfit, *On What Matters*.

[21] In principle, buck-passers who offer accounts of reasons in naturalistic or descriptive terms (e.g. Schroeder, *Slaves of the Passions*) are in a better position to advance this second strategy. As it stands though, this is a mere possibility. We know of no naturalistic reduction of reasons which would support the claim that incentives for attitudes are not reasons for them. For instance, Schroeder's view does not support the response condition (ibid., 165ff).

[22] Cf. D'Arms and Jacobson, "The Moralistic Fallacy"; Danielsson and Olson, "Brentano and the Buck-Passers"; Rabinowicz and Rønnow-Rassmussen, "The Strike of the Demon," 423; Schroeder, "Value and the Right Kind of Reason". Others who defend notions of fittingness on which it clearly won't be fitting to value the demon, without necessarily addressing such examples explicitly, include Brentano, *The Origin of our Knowledge*; Sidgwick, *The Methods of Ethics*; Broad, *Five Types of Ethical Theory*; Brandt, "Moral Valuation"; Ewing, *The Definition of Good*; McDowell, "Values and Secondary Qualities"; Chisholm, *Brentano and Intrinsic Value*; Wiggins, "A Sensible Subjectivism?"; Thomson, *Normativity*; Sigrún Svavarsdóttir, "Having Value and Being Worth Valuing," *Journal of Philosophy* 111 (2014): 84-109. We consider objections to this verdict below (3.1 and 3.2). For helpful further discussion see Richard Chappell, "Fittingness: The Sole Normative Primitive," *Philosophical Quarterly* 62 (2012): 684-704.

[23] Such a view is defended by, among many others, Paul Boghossian, "The Normativity of Content," *Philosophical Issues* 13 (2003): 31-45; Nishi Shah and David Velleman, "Doxastic Deliberation," *The Philosophical Review* 114 (2005): 497-534;Thomson, *Normativity*; Ralph Wedgwood, "The Aim of Belief," *Philosophical Perspectives* 16 (2002): 267-297. Others take belief to set knowledge as its standard of correctness, e.g. Declan Smithies, "The Normative Role of Knowledge," *Nous* 46 (2012), 265-88. This doesn't impugn our point here.

[24] Conor McHugh, "Fitting Belief," *Proceedings of the Aristotelian Society* 114 (2014):167-187; Svavarsdóttir, "Having Value".

[25] Alternatively, one could try to understand reasons directly in terms of value. For arguments against this 'value-first' view, see Jonathan Way, "Value and Reasons to Favour," in *Oxford Studies in Metaethics, Volume 8*, ed. Russ Shafer-Landau (Oxford: Oxford University Press, 2013) 27-49.

[26] The first of these views is suggested by Sharadin, "Reasons Wrong and Right", and Thomson, *Normativity*. The second is suggested by Chappell, "Fittingness: The Sole Normative Primitive". (Thomson and Sharadin use the term 'correct' rather than 'fitting', but we take them to have the same notion in mind).

[27] On reason as evidence, see e.g. John Broome, "Reply to Southwood, Kearns and Star, and Cullity," *Ethics* 119 (2008): 96-108; John Brunero, "Reasons and Evidence One Ought," *Ethics* 119 (2009): 538-45; Sean McKeever and Michael Ridge, "Elusive Reasons," in *Oxford Studies in Metaethics, Volume 7*, ed. Russ Shafer Landau (Oxford: Oxford University Press, 2012), 110-37. On reasons as explanations, see Stephen Kearns and Daniel Star, "Reasons: Explanations or Evidence?" *Ethics* 119 (2008): 31-56; John Brunero, "Reasons as Explanations," *Philosophical Studies* 165 (2013): 805-24.

[28] Jonathan Way, "Reasons as Premises of Good Reasoning," *Pacific Philosophical Quarterly* (forthcoming). See also Kieran Setiya, "What is a Reason to Act?" *Philosophical Studies* 167 (2013): 221-235.

[29] Other appropriate representations might include the perception that *p*. This would be one way to accommodate perceptual reasons within our account.

[30] See Way, "Reasons as Premises", for discussion and further precisions.

[31] In representing reasoning in this way, we follow John Broome, "Normative Requirements," *Ratio* 12 (1999): 398-419. Note that these are descriptions of reasoning, not derivations – it is not that having the first two attitudes entails having the third.

[32] On this alternative approach, it might still be extensionally correct to say that reasons are premises of good reasoning. Indeed, we think this claim is plausible, although we cannot defend it here.

[33] Plausibly, the point of reasoning is also to get rid of *un*fitting attitudes. For discussion of how to extend our account to reasoning that concludes in getting rid of or refraining from adopting an attitude, see McHugh, "Fitting Belief".

[34] This is not to deny that particular episodes of reasoning can have other specific aims. For example, you might reason about which attitudes would be beneficial, with the aim of acquiring them. In doing so, you would be reasoning towards an intention to acquire the beneficial attitude. All going well, this intention would be fitting. For more on the point of reasoning see Conor McHugh and Jonathan Way, "What is Good Reasoning?" (ms.).

[35] Clearly, a fully worked-out account would need to fill out this 'other things equal' clause. In other work (ibid.) we propose that a promising way to do this is in terms of *normality*. It is plausible that, given its function, testimony is true in normal cases; and similarly, given the function of promising, promises are normally permissible to keep.

[36] Broome, in his discussion of the view that reasons are evidence of what you ought to do ("Reply to Southwood, Kearns and Star, and Cullity"), discusses a similar example, which can be adapted for present purposes as follows:

<div style="text-align:center">

Belief that the reliable book says I ought to eat cabbage

</div>

leads to   Intention to eat cabbage.

According to Broome, the fact that the book says I ought to eat cabbage is not itself a reason to (intend to) eat cabbage – rather, it is (if anything) the facts about cabbage that give such reasons (cf. McKeever and Ridge, "Elusive Reasons"). If so, this is problematic for the view we have floated, since the transition in question is fittingness-preserving. However, if the moves we make below successfully deal with the case of admirable Jill, then they also deal with this case.

Note that there may be further relevant things to say about evaluative attitudes like admiration, that we lack the space to discuss here. For example, it may be that admiration is only ever fitting when held in some *respect*.

[37] For defence of the possibility of causation by the normative, see, e.g. Nomy Arpaly, *Merit, Meaning, and Human Bondage* (Princeton: Princeton University Press, 2006); Nicholas Sturgeon, "Moral Explanations," in *Essays on Moral Realism*, ed. Geoff Sayre-McCord (Ithaca: Cornell University Press, 1988); Wedgwood, *The Nature of Normativity*.

[38] It might be thought that this account is too restrictive. The fact that there were distinctive marks on the windowsill might be a reason to believe that Jones was the killer, even though the reasoning involves several steps. But our account can accommodate this. Often, complex patterns of reasoning are 'collapsible' into basic patterns with multiple premises. For example, from the premises that there were marks on the windowsill, that they were distinctive of Jones's shoes, etc., one can competently move *directly* to believing that Jones was the killer. By contrast, in the problem case this collapsing won't be possible, because the distinct steps of the reasoning do not simply correspond to further premises. The belief that Jill is admirable can't be packed in from

the start: it must be *based on* Jane's saying so, in order to exhibit sensitivity to the basic fittingness-preserving pattern. So the step to holding this belief is an ineliminable step in this reasoning, if it is to be done competently.

[39] More generally, we understand outweighing in terms of patterns of reasoning: roughly, where *p* is a reason to A and *q* is a reason not to A, *p* outweighs *q* if it is good reasoning to move from the beliefs that *p* and *q* to A-ing and not good reasoning to move from these beliefs to not-A-ing (cf. Way, "Reasons as Premises"). For defence of the idea that facts about the weights of reasons consist in facts about outweighing see Schroeder, *Slaves of the Passions*, Ch.7.

[40] Here we assume the supervenience of facts about what's fitting to value on non-evaluative facts. This plausible assumption is supported by Scanlon's own discussion.

[41] See Way, "Reasons as Premises".

[42] While plausible and popular, the response condition is not uncontentious; it is put under pressure by cases due to Elijah Millgram, "Williams' Argument Against External Reasons," *Nous* 30 (1996):197-220; Russ Shafer-Landau, *Moral Realism: A Defence* (Oxford: Oxford University Press, 2003); Schroeder, *Slaves of the Passions*. To use Schroeder's example (ibid, 165), we might think the fact that there is a surprise party for you at home is a reason for you to go home, although this is not a reason you could respond to. However, it is not obvious that we should accept this verdict – there are prospects for explaining the normative significance of the surprise party without accepting that it provides a reason to go home. Cf. Kieran Setiya, "Reply to Bratman and Smith," *Analysis* 69 (2009): 531-40; Neil Sinclair, "Promotionalism, Motivationalism, and Reasons to Perform Physically Impossible Actions," *Ethical Theory and Moral Practice* 15 (2012): 647-59.

[43] Such a view is also suggested by Broad, *Five Types of Ethical Theory*, 278; Danielsson and Olson, "Brentano and the Buck-Passers"; Chappell, "Fittingness: The Sole Normative Primitive".

[44] "The Strike of the Demon," 423.

[45] This might be resisted on the grounds that truth is a value. But if there is a plausible version of this idea, it is that true *belief* is valuable, not that the *object* of belief, a proposition, is valuable when true. In the case of valuing attitudes, by contrast, fittingness goes with the value of the object, not the attitude (cf. Way, "Value and Reasons to Favour").

[46] D'Arms and Jacobson, "The Moralistic Fallacy"; Christine Tappolet, "Values and Emotions: Neo-Sentimentalism's Prospects," in *Morality and the Emotions*, ed. Carla Bagnoli (Oxford: Oxford University Press, 2011) 117-34; Sharadin, "Reasons Wrong and Right".

[47] Schroeder, "Value and the Right Kind of Reason".

[48] Cf. W.D. Ross, *Foundations of Ethics* (Oxford: Clarendon Press, 1939), 278-9. The circularity worry arises on the assumption that pro-responses represent their objects as having normative properties like goodness. If this supposition is not made, then this 'accurate representation' account may not be a reduction of fittingness to something normative after all. So it would not be incompatible with our view.

[49] Shah and Velleman, "Doxastic Deliberation"; Daniel Whiting, "Should I Believe the Truth?" *dialectica* 61 (2010): 213-224.

[50] This is not to say that, in the wrong kind of reason cases, fittingness and permissibility *do* come apart. Indeed, the reader may notice that, since our view of reasons leads us to deny that incentives for attitudes are reasons for those attitudes, we are committed to denying that it is permissible to value the demon.

[51] For further arguments against taking correctness for belief to be permissibility, see Conor McHugh, "The Truth Norm of Belief," *Pacific Philosophical Quarterly* 93 (2012): 8-30, "Fitting Belief".

[52] In "A Sensible Subjectivism?" Wiggins offers an alternative story, which, while similar in spirit to ours, renders his account of value explicitly circular.

[53] Note that in taking for granted certain normative facts – e.g. that pleasure is fit to be desired – our view is at no disadvantage to its rivals. For instance, reasons-firsters typically take for granted facts about what sorts of facts provide reasons.

[54] Couldn't the resources we draw on here also be drawn on by the buck-passer, in order to give an account of the right-kind/wrong-kind distinction? In fact, Schroeder's view of right-kind reasons does draw on similar ideas. In "Value and the Right Kind of Reason" Schroeder claims that right-kind reasons with respect to A-ing are those reasons necessarily shared by everyone engaged in A-ing, because they are so engaged, along with all reasons which are derivative from such reasons. However, this account is problematic. First, it is not clear that it rules out wrong-kind reasons: if a necessarily existing god will reward all those who believe in it, this seems to give all believers a reason to believe in that god, because they are believers. Second, it has difficulty accommodating agent-relative right-kind reasons – for example, my reasons to intend to watch the game, because my team are playing - since these are not shared by all agents. Third, it seems a mistake to classify all reasons which derive from right-kind reasons as themselves right-kind reasons. For example, the fact that a false belief will generate many true ones is not a right-kind reason for that belief (cf. Way, "Transmission and the Wrong Kind of Reason"). Recently, Schroeder has retreated to the more non-committal claim that right-kind reasons for an attitude derive in some way from the nature of that attitude ("The Ubiquity of State-Given Reasons"). But the obvious way of spelling out this thought is to say that right-kind reasons are associated with

a standard of fittingness that is part of the nature of the attitude (cf. Sharadin, "Reasons Wrong and Right"). To give an account of value in terms of right-kind reasons so understood is to understand value ultimately in terms of fittingness, rather than reasons. On the assumption that fittingness is a normative property, such an account is thus incompatible with the reasons-first approach we are opposing here. Furthermore, even if such an account can be spelled out adequately, the buck-passer who wishes to avail of it must build it into her account of value from the start. By contrast, the fittingness-first account does not have to have such complications built in.

[55] Cf. Danielsson and Olson, "Brentano and the Buck-Passers"; Way, "Transmission and the Wrong Kind of Reason".

[56] Note that most existing accounts of right-kind reasons have this form – for instance, on the object-given/state-given theory (cf. n.15), right-kind reasons are reasons which depend on the object of the attitude. See Way, ibid., for further discussion.

[57] This is not to suggest that it's always easy to tell when an attitude is fitting. For example, it's not always obvious whether the moral considerations bearing on whether to respond in a certain way, such as the pettiness of envying a colleague or the offensiveness of being amused by a joke, make a difference to whether so responding is fitting. These cases do not cast doubt on the distinction between fittingness and other standards to which attitudes can be held. They just show that the normative question of whether a particular attitude is fitting is sometimes a difficult one. For discussion see Justin D'Arms and Daniel Jacobson, "Sentiment and Value", "The Moralistic Fallacy", "Wrong Kinds of Reason and the Opacity of Normative Force," in *Oxford Studies in Metaethics, Volume 9*, ed. Russ Shafer-Landau (Oxford: Oxford University Press, 2014) 215-44.

[58] This objection echoes an objection to value-first views of reasons made in Way, "Value and Reasons to Favour".

[59] "Value and the Right Kind of Reason", 29.

[60] For example, one might wonder how to generalise the explanation to cases where there are more than two respects of admirability. But this complication also arises for Schroeder's explanation in terms of the balance of reasons. What's crucial for present purposes is just that the degree of overall admiration that is fitting is a function of the degrees of admiration in various respects that are fitting.

Note that what we say here, combined with the linking principle and the account of the weights of reasons mentioned in n. 39, can also explain why, as Schroeder says, the fittingness of admiring Claire more than Mary is affected by the *reasons* to admire them, and by the weights of those reasons.