



# LUND UNIVERSITY

## The explanatory objection to the fitting attitude analysis of value

Orsi, Francesco; Garcia, Andrés G.

*Published in:*  
Philosophical Studies

*DOI:*  
[10.1007/s11098-020-01471-6](https://doi.org/10.1007/s11098-020-01471-6)

2021

*Document Version:*  
Peer reviewed version (aka post-print)

[Link to publication](#)

*Citation for published version (APA):*  
Orsi, F., & Garcia, A. G. (2021). The explanatory objection to the fitting attitude analysis of value. *Philosophical Studies*, 178(4), 1207-1221. <https://doi.org/10.1007/s11098-020-01471-6>

*Total number of authors:*  
2

*Creative Commons License:*  
CC BY-NC

### General rights

Unless other specific re-use rights are stated the following general rights apply:  
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

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

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

### Take down policy

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

LUND UNIVERSITY

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

---

# The explanatory objection to the fitting attitude analysis of value

Francesco Orsi<sup>1</sup>      Andres G. Garcia<sup>2,3</sup>

**Abstract:** The fitting attitude analysis of value states that for objects to have value is for them to be the fitting targets of attitudes. Good objects are the fitting targets of positive attitudes, while bad objects are the fitting targets of negative attitudes. The following paper presents an argument to the effect that value and the fittingness of attitudes differ in terms of their explanations. Whereas the fittingness of attitudes is explained, *inter alia*, by both the properties of attitudes and those of their fitting targets, the explanation of value tends to have a different content. In particular, objects have value in virtue of the features that make them valuable, and these need not involve any attitudinal properties. If this is right, then there are reasons to doubt the claim that for objects to have value is just for them to be the fitting targets of attitudes. Insofar as value is a property, it appears to be distinct from the property of objects being the fitting targets of attitudes.

**Keywords:** Fitting attitude analysis   Fittingness   Value   Attitudes   Normative explanation

## 1. Introduction

The fitting attitude analysis (FA) states that for objects to have value is for them to be the fitting targets of attitudes. Good objects are the fitting targets of positive attitudes, while bad objects are the fitting targets of negative attitudes. The

---

<sup>1</sup> University of Tartu, Tartu, Estonia

<sup>2</sup> Lund University, Lund, Sweden

<sup>3</sup> Humboldt University, Berlin, Germany

following paper presents an argument to the effect that value and the fittingness of attitudes differ in terms of their explanations.<sup>1</sup> Whereas the fittingness of attitudes is explained, *inter alia*, by both the properties of attitudes and those of their fitting targets, the explanation of value tends to have a different content. In particular, objects have value in virtue of the features that make them valuable, and these need not involve any attitudinal properties. If this is right, then there are reasons to doubt the claim that for objects to have value is just for them to be the fitting targets of attitudes. Insofar as value is a property, it appears to be distinct from the property of objects being the fitting targets of attitudes.

The next section includes a formulation of the explanatory objection and distinguishes our objection from a different worry about FA with which it might be mistaken. The third section defends the claim that the properties of fitting attitudes help explain why they are fitting towards their targets, while the fourth section defends the claim that the properties of fitting attitudes are not as a rule needed to explain value. Each of these two claims is supported by providing a brief intuitive motivation and by replying to various objections. In the end, the paper illustrates that if FA is to be reconciled with the apparent explanatory differences between value and the fittingness of attitudes, then this will require fresh work on part of its advocates.

## 2. The explanatory objection

In what follows, we take FA to be the claim that for objects to have value is for them to be the fitting targets of attitudes.<sup>2</sup> There are several ways to understand FA. One is as a claim of conceptual identity: value concepts are identical to concepts of fitting attitudes. Another is as a claim of metaphysical identity between properties (whether or not conceptual identity holds): for example, the property of being good (or having positive value) is identical to the property of being the fitting target of positive attitudes, and the property of being bad (or having negative value) is identical to the property of being the fitting target of negative attitudes, and so on for other value properties. A third way to understand FA is as a claim of metaphysical reduction: the property of being good, for example, is nothing over and above the property of being the fitting target of positive attitudes. Property P reduces to property Q when property Q exhaustively explains property P. In this sense, reduction is, unlike identity, an asymmetrical relation. But in what follows the difference between identity and reduction will not matter. The explanatory objection

---

1 I.e. in terms of their “normative explanation” (Väyrynen 2013)—an explanation citing the properties in virtue or because of which something is good, right, fitting etc. We take no stand on how normative explanation relates to notions of supervenience or grounding.

2 Our focus in this article is on the FA analysis of thin value properties. In Sect. 4 we consider a point related to the thin/thick distinction.

that is about to be presented is meant to cast doubt on both the metaphysical identity and the reduction versions of FA.<sup>3</sup> Here is the argument:

- P1 If property P is identical with or reducible to the property Q, then any fact that explains why Q is instantiated also explains why P is instantiated.<sup>4</sup>
- P2 Not all facts that explain why the property of being a fitting target of attitudes is instantiated also explain why the property of having value is instantiated.
- C Therefore, the property of having value is neither identical with nor reducible to the property of being a fitting target of attitudes.

P1 is not the strong claim that P cannot be identical with or reduced to Q unless both properties share the exact same explanatory base. P1 is the weak claim that P cannot be identical with or reduced to Q in the event that there is some fact that explains Q but that does not also explain P. While the strong claim represents a tempting application of the principle of identity of indiscernibles to properties, it is not really needed for the explanatory objection to work. It is also important to note that the relevant notion of explanation that is invoked by both premises is meant to be metaphysical and not just linguistic or epistemic, meaning that the explanatory base of a property is here understood in terms of the facts that metaphysically explain its instantiation. The assumption is that a property cannot be identified with or reduced to another property unless the latter traffics in the same explanatory relations as the former.

P1 is here applied to value and the fittingness of attitudes, but it can be illustrated by examples taken from a variety of domains and explanatory contexts. If it is discovered that the facts that explain why a person has the property of being in a certain brain state do not also explain why that person has the property of being happy, then it is reasonable to suggest that the latter property is neither identical with nor reducible to the former. Similarly, if it is discovered that the facts that explain why an apple has the property of reflecting light at a wavelength of 625–740 nm do not also explain why that apple has the property of being red, then it is reasonable to suggest that the latter property is neither identical with nor reducible to the former.<sup>5</sup>

P1 is here assumed to be the least controversial of the premises and so most of the following discussion focuses on supporting P2. This is achieved by defending the following claims:

---

<sup>3</sup> In line with most of the literature, we understand FA as stronger than the claim that there is merely a necessary coinstantiation of value and fitting attitudes. For the sake of simplicity and concision, we also deliberately choose the fitting attitude analysis as our target rather than the buck-passing account of value (Scanlon 1998; Rowland 2019), which relies on the notion of reasons for attitudes. While we believe that many of the points we will be making can be applied to the buck-passing account as well, cashing this out in detail would require a longer investigation into the nature of reasons that would take us beyond the scope of the paper. See also Jacobson (2011) on differences between FA and the buck-passing account.

<sup>4</sup> We are assuming that a reduction of one property to another does not entail that there is an identity between them.

<sup>5</sup> If these examples fail to convince, we can reformulate premise 1 in terms of best explanation: If a fact explains the instantiation of Q, but fails to explain the instantiation of P, then a view on which P is not identical with nor reducible to Q is pro tanto more plausible than a view on which P is identical with or reducible to Q.

1. Fittingness needs attitudes The fittingness of attitudes is always explained, *inter alia*, by both properties of the fitting attitudes A and properties of their fitting targets O.
2. Value does not need attitudes Value is explained, *inter alia*, by the properties that make objects valuable, and while these need to include the properties of the object O, they need not include properties of the attitude A.<sup>6</sup>

If claims (1) and (2) are correct, then P2 follows: Not all facts that explain why the property of being a fitting target of attitudes is instantiated also explain why the property of having value is instantiated. The explanation of the fittingness of attitudes is importantly different from the explanation of value, and this precludes—or at least puts at a serious disadvantage—the thesis that value is identical with or reducible to the fittingness of attitudes. FA advocates thus need to show that at least one of the two claims is false, i.e. they need to show that the explanation of fittingness and the explanation of value are not so different after all. This could be achieved by showing, *contra* (1), that the explanation of fittingness does not need attitudinal properties, or by showing, *contra* (2), that the properties that make objects valuable need to include properties of the attitude A, or again by showing that FA can actually accommodate both claims (1) and (2). We argue below that (1) and (2) are plausible claims that any theory hoping to capture the relation between value and the fittingness of attitudes should strive to accommodate. And we argue in Sect. 4 that a version of FA that tries to accommodate both claims (1) and (2) (rather than reject either) is highly problematic.

Before moving on to the defence of claims (1) and (2), it should also be noted that the explanatory objection to FA does not rely on any specific views regarding the formal structure of the properties under consideration. For example, the explanatory differences between value and the fittingness of attitudes will not be established here by the observation that while value appears to be a monadic (one-place) property of an object or of a state of affairs, the fittingness of attitudes appears to be a polyadic (many-place) one—fittingness appears to be (at least) a two-place relation between an attitude and an object (Dancy 2000). While it is tempting to conclude that such a difference in formal structure entails an explanatory difference, this inference is unjustified. Nothing in the nature of a polyadic property (or a relation) requires that all terms of the relation contribute in explanatory terms: thus, even if fittingness were a two-place relation between an attitude and an object, properties of the attitude may fail to contribute in explanatory terms (compare: unicorns arguably fail to contribute to explaining the relation Lucy

---

<sup>6</sup> Rabinowicz and Rønnow-Rasmussen (2004) have pointed out that the properties of fitting attitudes A inevitably give rise to corresponding properties in their fitting targets O. For the sake of illustration, suppose artistic achievements are the fitting targets of admiration and that admiration has the property of being V. This entails that artistic achievements have the property of being such that they are the fitting targets of attitudes that are V. While this may seem to complicate matters somewhat, it does not constitute a big problem for the explanatory objection or the claims just made. It just needs to be remembered that whenever the properties of O are addressed in what follows, we have in mind specifically the properties of O that are not ultimately explained by the properties of A.

loves unicorns). Likewise, even if the property of being good is monadic, the explanation of why something is good might still include facts about things other than the value bearer, such as facts about attitudes (compare: 2 being a prime number, though formally monadic, can only be explained by reference to other numbers, namely number 1). In other words, our claims (1) and (2) require an investigation into the respective explanations of fittingness and value.

### 3. Fittingness needs attitudes

The following section focuses on defending claim (1), according to which the fittingness of attitudes holds, *inter alia*, in virtue of both the properties of those attitudes (what might be called their “shape”) and the properties of their fitting targets. The claim appears to have as much immediate intuitive appeal as the observation that the fittingness of a glove to a hand (or a key to a lock) is explained by the properties of the glove (key) as well as by the properties of the hand (lock). However, more needs to be said in support of claim (1) than appeals to analogies with the notion of fittingness as it applies to physical objects. We believe that additional support for the claim can be derived from considerations having to do with how different value bearers end up being the fitting targets of different types of attitudes. Two examples will be developed below focusing on artistic achievements and the virtue of kindness, respectively.

Artistic achievements are the fitting targets of admiration, but not of anger, and this is partly because of the properties that make things into artistic achievements: their originality, aesthetic qualities, and so on.<sup>7</sup> The question remains why objects with these particular properties are matched by the attitude of admiration and not by the attitude of anger. A plausible answer to this question must invoke the properties that make admiration different from anger. We believe that we do not need to take any definitive stand on the question what these properties are, as this is as much of a substantive issue as the one about what properties make objects valuable. However, in the interest of clarity, we will nevertheless mention some of the more obvious possibilities of what the properties in question might be. Admiration could be afforded the right shape by its functional properties and connection to human behaviour. It is arguably in the nature of admiration that any person that admires an object will be disposed to protect and preserve that object, whereas it is in the nature of anger that any person who is angry at an object will be disposed to do the opposite of these things. What makes admiration a fitting response to artistic achievements could also be a matter of its intentional character. The idea is that admiration presents its targets as having the kinds of properties that artistic achievements have, whereas anger presents its targets as having the kinds of

---

<sup>7</sup> Note that there may be cases where it is fitting to be angry at an artistic achievement, but the point is that it will not be in virtue of the properties that make it an artistic achievement. The question discussed here is what attitudes it is fitting to direct toward artistic achievements as such.

properties that artistic achievements lack (D'Arms and Jacobson 2000).<sup>8</sup> These suggestions seem to resonate well with the claim that attitudes have inherent standards, and that artistic achievements answer to the standards of admiration but not of anger (McHugh and Way 2016).

Consider also the virtue of kindness and how it tends to make kind people fitting targets of respect but not of amusement. It seems clear that an explanation for this must invoke the traits that make people kind: their empathy, willingness to help others, and so on.<sup>9</sup> However, this is not enough to explain why people with these particular traits are the fitting targets of respect and not amusement. Just as with the previous example, it also seems necessary to invoke the qualitative differences between the relevant attitudes. To reiterate, we do not need to provide any particular view of what these properties are, since that requires engagement in substantive debates that would take us far beyond any general discussions about how value relates to the fittingness of attitudes. Nevertheless, it is worth mentioning that the possibilities broached in the previous paragraph seem just as applicable here. The attitude of respect could be afforded the right shape by its functional properties and connection to human behaviour. It is arguably in the nature of respect that any person that respects another will be disposed to emulate that other person, whereas it is in the nature of amusement that any person who is amused by another will be disposed to laugh at that other person. Just as before, what makes respect a fitting response to kind people could also be a matter of how it presents the properties of its targets. Respect presents its targets as having the kinds of properties that kind people have, whereas amusement presents its targets as having the kinds of properties that amusing people have. These suggestions also illustrate a sense in which kind people as such meet the standard inherent in the attitude of respect, while kind people as such fail to meet whatever standard is inherent in the attitude of amusement.

What these examples show is that two complementary questions need to be answered in order to explain why it is fitting to direct attitudes toward certain objects. The first asks what it is about the objects themselves that make them the fitting targets of the attitudes, while the second asks what it is about the attitudes that make them a fitting response to the objects. An explanation that only answers one of these questions cannot hope to offer a complete account of the fittingness of attitudes.

Besides referring to substantive examples to support this view, it could also be pointed out that some FA advocates are already committed to it. For example, Daniel Jacobson puts himself among those FA advocates when he writes:

Maybe pleasure, knowledge, beauty, and friendship are all desirable or admirable things, where this means not that we can desire or admire them but

---

<sup>8</sup> Another example involves a person's own artistic achievements, which may be the fitting targets of pride, while the artistic achievements of complete strangers are not. The nature of pride as a typically self-regarding attitude helps explain the difference in fit.

<sup>9</sup> Cases can of course be imagined where it is fitting to be amused at kind people, but as before, it will not be in virtue of the properties that make people kind. It is important to remember that the question discussed here is what attitudes it is fitting to direct toward kind people as such.

that we should (ought, have reason to) do so, in virtue of facts about human nature, the inherent qualities of those attitudes, and natural facts about the objects (2011; italics added).

When he talks of the “inherent qualities” of desire and admiration, we take it that Jacobson has in mind the shape of those attitudes. It is interesting that in addition to the shape of attitudes and the properties of their targets, Jacobson also refers to facts about human nature as being explanatorily relevant. It is unclear to us what these facts might be, but him being right about this would of course only strengthen the explanatory objection. For while facts about human nature may help explain the fittingness of attitudes, intuition suggests that facts about human nature need not be part of the explanation of value. We hereafter leave the potential relevance of facts about human nature aside and continue to focus on claims (1) and (2).

Jacobson is explicit in endorsing claim (1), but other FA advocates commit themselves to it implicitly because of their treatment of the wrong kind of reasons problem.<sup>10</sup> The problem is that there seem to be cases where value and the fittingness of attitudes come apart. The usual example involves the case of an evil demon who threatens people with torture unless they admire him for his own sake. This seems to be a case where it is fitting to admire the demon for his own sake, but this would commit FA to the demon having positive value. Some FA advocates have replied by dissolving the problem: while it might be fitting to have some other response like bringing it about that one admires the demon for his own sake, it is just not fitting to admire the demon for his sake, and thus it does not follow that the demon has positive value.<sup>11</sup> Now, it seems that the reason why an attitude like admiration is not fitting towards the evil demon has partly to do with the nature or shape of admiration, and not just with the nature of the demon. Certain properties of admiration make it an unfitting response to the evil demon, regardless of how threatening he is.<sup>12</sup> If this is right, then it is natural to assume that in those instances where it is fitting to admire something or someone, the shape of admiration must also be part of the explanation why.

To remind the reader of the dialectic here: if some FA advocates turn out to be explicitly or implicitly committed to our claim (1) (Fittingness needs attitudes), then in the context of our argument their task will be to show that our claim (2) (Value does not need attitude) is false—or that it can somehow be reconciled with claim (1).

We can identify two major reasons for doubting claim (1). The first one hinges on the idea that the fittingness of certain attitudes may lack an explanation altogether. A fortiori these would be cases where the explanation of fittingness does not require

---

10 See Rabinowicz and Rønnow-Rasmussen (2004) and Gertken and Kiesewetter (2017) for excellent discussions.

11 Proponents of the strategy just described include Parfit (2011: App. A), Skorupski (2007), Rowland (2014), and Way (2012).

12 This kind of solution is considered, though rejected, in Rabinowicz and Rønnow-Rasmussen (2004: 422–423). It is defended by Danielsson and Olson (2007), albeit in terms of correctness rather than fittingness, and by McHugh and Way (2016).



the shape of attitudes. For example, it might be argued that the fittingness of having a negative attitude toward undeserved pain is primitive in this sense.<sup>13</sup>

The objection seems to us to overshoot its mark, for it is plausible to assume that normative facts hold in virtue of other facts.<sup>14</sup> The most obvious fittingness facts are no exception. If it is fitting to have a negative attitude towards undeserved pain, this is a normative fact that holds in virtue of other facts, presumably facts about the nature of undeserved pain and, as we argue here, also facts about the shape of negative attitudes. To claim otherwise seems to us as strange as the claim that objects can have value without there being something about the objects making them valuable.<sup>15</sup> Moreover, even if there were cases where the fittingness of attitudes is in fact primitive in the sense of lacking an explanation, these cases would not affect claim (1). Our point remains: when fittingness facts admit of an explanation, these facts will be explained, *inter alia*, by the shape of the fitting attitudes.

Finally, if it is true that some fittingness facts can lack an explanation, but value facts are always explained by some value-making property, this would support rather than weaken our explanatory objection to FA. It is difficult to understand how a fact that is always in need of a metaphysical explanation could possibly be identified with or reduced to another fact that does not need such an explanation. We therefore suggest that all putative cases where the fittingness of attitudes lacks an explanation actually speak against the proposed identity or reduction as well.

The second major reason for doubting claim (1) hinges on the claim that all the explanatory work that is apparently done by the properties of fitting attitudes must be done by the properties of their fitting targets.<sup>16</sup> Rowland may seem to express this idea when he maintains that “it is the nature of pleasure that both grounds the reason to have pro-attitudes towards pleasure and makes it valuable” (change “grounds the reason” with “makes it fitting”) (2019: 167 fn. 20). And McHugh and Way state: “the total nonevaluative facts [about a thing] are sufficient for goodness—for it to be fitting to value the thing” (2016: 594). To be sure, these claims are usually made for the purposes of clarifying that, on FA or its close relative, the buck-passing account of value, evaluative properties such as goodness do not explain why an attitude is fitting to O or why there is reason to have a pro-attitude to O. As in

---

<sup>13</sup> This objection was put to us by an anonymous referee.

<sup>14</sup> Whether these other facts are only non-normative is not important for our point here.

<sup>15</sup> If the complaint is that the explanation throws no real light on the fittingness of having a negative attitude toward undeserved pain, then our reply is just a reminder that the notion of explanation that is used in this context refers to a genuine metaphysical relation that does not necessarily have any epistemic benefit. The fact of the matter is that the shape of attitudes is often taken as a given and therefore rarely mentioned explicitly. It might even seem bizarre in most cases to ask why a negative response is the fitting one in the face of undeserved pain, but this is just because people are expected to know about the nature and relevance of the negative response. What seems more interesting from an epistemic standpoint are the other factors that play a role for the response in question, but of course, this does nothing to show that attitudinal properties are metaphysically irrelevant.

<sup>16</sup> Where these properties do not include or mention facts about the very fitting attitudes, although of course they might include facts about attitudes, if the valuable object is itself an attitude, e.g. love for knowledge.

Scanlon's famous metaphor, the "buck" (the job of providing reasons for attitudes, or making attitudes fitting) is passed from goodness to whatever makes the object good. It is important to note, then, that the statements by Rowland and McHugh and Way are not meant as rejections of our claim (1). In fact, we do not know of any detailed discussion ever dedicated to claim (1), despite the amount of work currently being done on the fittingness of attitudes.<sup>17</sup> What is certain, though, is that both Rowland's and McHugh and Way's claims display sensitivity to the explanatory concerns voiced in our article. In other words, whether as buck-passers or FA advocates, they seem to assume that whatever explains the relevant reason/fittingness facts must also explain value facts, and vice versa, because the truth of their views hinges on this. Hence the way they seamlessly move back and forth from "grounds the reason to have a pro-attitude towards x/is sufficient for it to be fitting to value x" to "makes x valuable/is sufficient for x's goodness."

So, regardless now of Rowland's or McHugh & Way's intentions, how could one argue that the shape of attitudes contributes nothing to their fittingness towards an object over and above the contribution given by the object's properties? Such an argument could start from this apparently plausible claim about the connection between fittingness and reasons: "any fact that makes an attitude fitting provides a reason for that attitude" (Howard 2018: 3). If this claim is true, then it would seem to follow that the shape of an attitude cannot even partly make an attitude fitting, if the shape of an attitude cannot provide a reason for that attitude:

- (i) If fact *f* makes an attitude fitting, then *f* provides a reason for that attitude.
- (ii) The shape of the attitude cannot provide a reason for that attitude.
- (iii) Therefore, the shape of the attitude does not make an attitude fitting.

We believe premise (i) is false. Here is a ready counterexample: the truth of *p* seems to be part of what makes it fitting to believe that *p*. But the truth of *p* does not, normally, provide a reason for believing that *p*.<sup>18</sup> Reasons for belief that *p* are normally provided by facts which are relevantly related to *p*, in that they provide evidence for *p*. But *p* cannot provide evidence for itself. Therefore the truth of *p* cannot be a reason for believing that *p*. Or, at any rate, it is a very controversial question whether the truth of *p* provides a reason for believing that *p*, whereas it seems much less controversial that the truth of *p* makes it fitting to believe that *p*. Denying the latter is dangerously close to denying the intuitive thought that there is at least something correct or fitting about a belief that matches the truth or what is the case.<sup>19</sup> In short, this is a case where, contrary to premise (i), whether a given fact makes an attitude fitting does not depend on whether it provides a reason for that attitude. So, should it turn out that the shape of an attitude does not provide a reason for that attitude, it would not follow that the shape of an attitude fails to contribute to making the attitude fitting.

---

<sup>17</sup> See the otherwise very informative (Howard 2018).

<sup>18</sup> It might be said that *p* counts in favour of believing that *p*. But this truism is better expressed as a claim about *p* making it fitting to believe that *p*—and this is so partly, as our story goes, because of the shape of belief.

<sup>19</sup> We thank a reviewer for inviting us to expand on this.

It might be said that rejecting premise (i) has a high price: if “fitting-makers” for attitudes are not guaranteed to provide reasons for the attitudes, it is not clear in what sense facts about fittingness can be said to be normative facts. However, we think this worry is misplaced. It is commonly accepted that there is a difference, at least in principle, between for example reasons why an action is right or what you ought to do, and reasons for you to perform that action. The former are explanatory reasons of some sort, the latter are practical. Not all of the former need to be included in the latter. But facts about rightness or oughts are not made any less normative just because some facts which explain why an action is right may fail to provide reasons for a given agent to perform that action. By the same token, if there are facts that explain why an attitude is fitting, but do not provide reasons for that attitude, this does not make facts about fittingness non-normative facts. It is just that not everything that makes an attitude fitting needs to have a “second life” as a provider of a reason for that attitude. Facts about fittingness still retain their normativity insofar as it is generally true that if it is fitting to have attitude A, then there are reasons to have that attitude. Moreover, the facts that make attitude A fitting without providing a reason for it are still normatively relevant, because if they were not the case, attitude A would not be fitting, and therefore there would not be a certain kind of reason to have that attitude (there might be other reasons to have that attitude, unrelated to the fittingness of the attitude, but the point holds regardless).<sup>20</sup>

Thus we remain unconvinced by attempts to show that the explanation of fittingness does not include the shape of attitudes.

#### 4. Value does not need attitudes

The previous section focused on defending claim (1), according to which the fittingness of attitudes holds, *inter alia*, in virtue of both the properties of those attitudes (what might be called their “shape”) and the properties of their fitting targets. This section is instead dedicated to defending claim (2), according to which value is explained, *inter alia*, by the properties that make objects valuable, and while these need to include the properties of value bearers, they need not include the shape of the attitudes that would be fitting towards the value bearer. Recall that once claims (1) and (2) are both established, the second premise of the explanatory objection P2 is also established: Not all facts that explain why the property of being a fitting target of attitudes is instantiated also explain why the property of having value is instantiated. The conclusion that value cannot be identified with or reduced to the fittingness of attitudes looms.

Claim (2) is supported by Moorean intuitions about the attitude-independence of value. Imagine a pair of worlds of which one is more beautiful than the other (if beauty is an unconvincing example, replace it with fairness). If all else is equal, then the fact that one world is more beautiful (or more fair) than the other seems

<sup>20</sup> We prefer to remain neutral, for reasons of space, about whether premise (ii) is true.

sufficient to explain why it is also better than the other. The properties of a fitting attitude—e.g. the properties of preference—do not need to be invoked in order to explain why the first world is better. Even if the first world is fittingly preferred, or is a fitting target of a preference over the other, facts about the shape of preference are superfluous at best and irrelevant at worst to the metaphysical explanation of why one world is better than the other. In this sense, value facts are cheaper than fittingness facts: their explanation does not require the shape of attitudes.

Consider also a substantive axiology according to which pain is intrinsically bad. It is unclear why this substantive axiology should be made to include a story about the shape of fitting attitudes in addition to the intrinsic nature of pain. It is also unclear what attitudes could even be relevant to the case at hand, for while certain attitudes are clearly fitting towards pain—say, attitudes of disfavour or aversion—it seems very strange to suggest that the shape of those attitudes explains why pain is intrinsically bad. It is strange not because the relevance of attitudinal properties is already known about and hence unnecessary to mention, but because such properties seem altogether irrelevant.

We readily acknowledge that the shape of a fitting attitude *A* towards an object *O* plays an explanatory role when attitude *A* is part of the very valuable object *O*. For example, the desire for justice may be something valuable on top of justice being valuable, and it may be true that it is fitting to desire that people have a desire for justice. The attitude of desire in this case is both fitting towards *O* and part of *O* itself. In this case, it is plausible that the shape of desire contributes to explaining why a desire for justice is something valuable on top of justice being valuable. But these are somewhat special cases—clearly it would be an inadmissible constraint to require that the shape of a fitting attitude towards any object be always part of the value-making properties of that object. Such a constraint would leave room only for a very peculiar axiology, where only a specific kind of states of affairs are valuable: those which include the very attitudes that are fitting towards them. We thus take it that our claim

(2) (Value does not need attitudes) is not only intuitively plausible, but that rejecting it has implausible consequences. FA theorists, like anybody else, had better accept it.<sup>21</sup>

But if FA theorists accept claim (2), then they are faced with a choice: they must either reject our claim (1) (Fittingness needs attitudes), or find a way to accept both (1) and (2) and show that the combination of (1) and (2) is consistent with value and fittingness having the same explanatory bases, and thus consistent with value facts being identical or reducible to fittingness facts. In the previous section we argued that rejecting claim (1) is implausible. In the rest of this section we consider the prospects for the second option. Can FA embrace both Fittingness needs attitudes and Value does not need attitudes?

---

<sup>21</sup> Note that Value does not need attitudes does not beg the question against FA. FA as such does not require that value be partly explained by the shape of fitting attitudes. As remarked in the text, there is an alternative for FA: accepting (2) and rejecting (1), hence maintaining that neither value nor the fittingness of attitudes are partly explained by attitudinal properties. This is a problematic but genuine alternative, which means that it is not a necessary claim of FA that value be explained by the shape of attitudes. For example, rejection of (1) is explicit in Olson (2004), who rules out any reference to the fitting attitude as part of the explanation why it is fitting to take that attitude towards a given valuable object. Rejection of

(1) might also be implicit in Rowland's quote reported above (see also Ewing 1948: 172).

In principle, it could. Note how we formulated Value does not need attitudes:

Value is explained, *inter alia*, by the properties that make objects valuable, and while these need to include the properties of value bearers, they need not include any properties of the attitudes that would be fitting towards them.

FA advocates might agree with us that the properties that make objects valuable need not include any attitudinal properties. But then they might argue that properties of attitudes are a necessary part of the explanation of value facts as part of the *alia* which our formulation leaves room for. The shape of attitudes may have a systematic role to play in the complete explanation of value, even though not necessarily *qua* the properties that make objects valuable. In this way, the shape of attitudes would play an explanatory role with respect to both fittingness and value facts, and any apparent explanatory difference between fittingness facts and value facts would disappear.

While this is a move that, to our knowledge, no FA advocate has ever spelled out, the idea might be that the shape of an attitude plays the role of explaining why other facts or properties make an object valuable, without itself making the object valuable.<sup>22</sup> The most promising model here seems to be Jonathan Dancy's notion of enabling conditions (Dancy 2004). For example, on some views the fact that a certain pleasure is morally permissible enables the fact that I am feeling pleasure to make my experience *pro tanto* good, without thereby adding to the facts that make my experience *pro tanto* good. FA advocates may likewise say that the shape of an attitude enables other facts to play the value-making role; for example, the shape of admiration may enable other facts, e.g. facts about the artistic qualities of a painting, to make a given painting admirable. In this way the shape of admiration (or any other attitude) would contribute to explaining why other facts make an object valuable.

This is a theoretically available strategy. However, we find it deeply problematic. This move would violate the substantive neutrality that is understood to be a *sine qua non* of FA and of players to this debate in general.<sup>23</sup> As with the example of morally permissible pleasure, it is (usually) a substantive thesis to claim that a given fact or property plays an enabling role with respect to other facts or properties. Not everyone needs to agree that a pleasure is good only if morally permissible; likewise, not everyone (defending or not FA) needs to agree that certain properties are value-making only if attitudes have the shape that they do. By including properties of attitudes in the explanation of value, FA would be wedded to a particular view in substantive axiology, albeit one about specific enabling conditions rather than about specific value-making properties. This would be a significant theoretical cost to bear. (Note that our own claim that the fittingness of attitudes is partly explained by the shape of attitudes is safe from the charge of non-neutrality: the claim that the shape of an attitude partly explains why that attitude is fitting towards a given object is no more substantive than the claim that certain

---

<sup>22</sup> We thank a reviewer for pressing this point.

<sup>23</sup> See Rowland (2019).

properties of an action partly explain why that action is right. Everyone in normative ethics can accept the latter claim, and likewise everyone in the theory of fitting attitudes can accept our claim.)<sup>24</sup>

The FA advocate may reply that understanding shapes of attitudes as enablers for other properties to be value-makers is a substantive view, but still neutral enough. After all, the idea that the shape of admiration explains why certain other properties—whatever they may be—make an object admirable would seem to be compatible with any first-order view about what is admirable or about what makes something admirable. While we believe this claim to be questionable (and anyway in need of thorough examination), we have something more to add in our own defence even if we were to grant this claim. Suppose it were true that the full explanation of something being admirable, enviable, fearsome etc. owed something to the shape of admiration, envy, fear etc. It would still need to be shown that the full explanation of the rather different property of being good (bad, better etc.) simpliciter owes something to the shape of attitudes.

At this point the proponent of this strategy has two options, and both are unpalatable. For one, she could argue that there is no such thing as value simpliciter—there are only “response-mentioning” value properties (admirability, enviableness, fearsomeness etc.). However, whatever the intrinsic merits of this view, it seems too high a cost to bear for FA to be forced to adopt it in a defence from our objection. One would have hoped FA to avoid having to take a stand on what kinds of value properties there are.

The second, less radical, option is to take a modest view of the scope of FA, and claim that FA should only be an account of such “response-mentioning” value properties (and perhaps of other value properties lying rather on the right-hand side of the thin/thick distinction, such as being good for someone and being good as a kind), while leaving it open whether FA can also account for value simpliciter and other “thinner” value properties. This is in fact a view hinted at by Jacobson:<sup>25</sup>

One might hold an FA theory specifically for those values that seem response-dependent, while thinking that certain other values—including perhaps the good or moral worth—are autonomous from human response and attitudes. (I take this complex view to have considerable plausibility.) (2011)

Our reaction to this move is that such relative plausibility comes at the price of attractiveness. One would have hoped FA to be ambitious enough to account for the kinds of value properties that have been the subject of traditional axiological debates. Value simpliciter is the property that hedonists ascribe to pleasure,

---

24 A similar worry would apply, were FA to borrow from the toolkit of contemporary constructivism and subjectivism, by claiming that rather than playing the role of value-makers or enablers, the properties of attitudes are part of the constitutive grounds of value. While this may help FA advocates avoid the explanatory objection, it seems fair to point out that a version of FA that does not need to borrow from the toolkit of particular meta-ethical views (such as constructivism and subjectivism) would be preferable, all else being equal. See Rønnow-Rasmussen (2011, ch.1) and Fritzson (2014) for discussions of constitutive grounds.

25 That is probably why he is among the few to explicitly mention the “inherent qualities” of attitudes as part of the explanation of why attitudes are fitting.

preferentists to preference satisfaction, Kantians (arguably) to the good will, objective list theorists to a number of things like knowledge, friendship, beauty, and so on.<sup>26</sup> Moreover, value simpliciter does seem *prima facie* to be a subject of ordinary, and not just philosophical, evaluative thought and talk. Any analysis of value that eschews the project of analysing value simpliciter must therefore be to that extent less attractive. Finally, if FA were just the thesis that for *x* to be admirable (desirable etc.) is for *x* to be the fitting target of admiration (desire etc.), FA would amount to little more than a philosophical (metaphysical or conceptual) articulation of a semantic platitude, for in English “admirable” quite simply means “a fitting (or cognate terms: appropriate, worthy etc.) target of admiration”, and the same *mutatis mutandis* for other response-mentioning value terms.<sup>27</sup> There is hardly any lexical room for anyone to disagree with a version of FA restricted to such value properties. There would of course remain important philosophical work to be done on the nature of fittingness, and on what makes an attitude fitting towards certain objects, but this is something that goes beyond the analysis of value as such: one can investigate the nature of fitting attitudes regardless of endorsing FA.

In conclusion, we see little prospect for a plausible and attractive version of FA that is capable of accommodating both our claims (1) and (2).

## 5. Concluding remarks

In this paper, we have laid out a novel argument against FA understood as the claim that value facts are identical or reduce to facts about fitting attitudes. The fittingness of attitudes holds, *inter alia*, in virtue of both the shape of fitting attitudes and the properties of their fitting targets, whereas the explanation of value need not involve the shape of the fitting attitudes. If this is true, then having value can hardly be identical with or reducible to being the fitting target of an attitude.

If our argument is on the right track, this leaves FA advocates with two options: either bite the bullet of rejecting our claims (given the many benefits that FA may otherwise have over all rivals), or come up with a formulation of FA that is immune to our explanatory concerns. We take it that choosing the former option would be surprising, to say the least: certainly no FA advocate is willing to bite the bullet of allowing wrong kinds of reasons for admiration to determine what is admirable, or that of allowing that solitary goods and evils are not really good or bad states of affairs.<sup>28</sup> Since there seems to be no more reason to bite the bullet in the case of our explanatory objection than there is in those cases, the preferable option for FA is to do fresh work in order to avoid our objection.

<sup>26</sup> We are not making the bold claim that axiology must talk about value simpliciter. We do recognize that some of these views are put forward, for example, only as theories of well-being.

<sup>27</sup> Of course we do not mean to deny that these terms in English may also have purely descriptive, non-evaluative meanings, e.g. “such as to arouse admiration (envy, fear etc.)”.

<sup>28</sup> On the problem of solitary goods (intuitively good states of affairs that, apparently, it is not fitting or there is no reason for anyone to favour), see Dancy (2000), Bykvist (2009), Orsi (2013), Reisner (2015).

Acknowledgements: Both authors are very grateful to Wlodek Rabinowicz and Toni Rønnow-Rasmussen for commenting on an early draft of the paper. We are also grateful to two reviewers of this journal for their extremely helpful comments. Francesco Orsi thanks audiences in Tartu, Moscow (HSE) and Geneva (Thumos Seminar) for their feedback. Orsi's contribution was supported by Eesti Teadusagentuur (Grant No. PUT1630). Andre's G. Garcia presented this paper at the colloquium for practical philosophy at Humboldt-Universität zu Berlin and at the Centre for Aesthetic, Moral, and Political Philosophy at Leeds University. Gratitude is owed to all the participants of these events for their suggestions as well, especially Tomas Schmidt, Razvan Sofroni, Philip Fox, Jack Woods, Pekka Väyrynen, Gerald Lang, and Alex Jackson. Garcia's contribution was supported by the Swedish Research Council (Grant No. 2018-06612).

## References

- Bykvist, Krister. (2009). No good fit: Why the fitting attitude analysis of value fails. *Mind*, 118(469), 1–30.
- D'Arms, Justin, & Jacobson, Daniel. (2000). The moralistic fallacy: On the 'appropriateness' of emotions. *Philosophical and Phenomenological Research*, 61(1), 65–90.
- Dancy, J. (2000). Should we pass the buck? *Royal Institute of Philosophy Supplements*, 47, 159–173.
- Dancy, J. (2004). *Ethics without principles*. Oxford: OUP.
- Danielsson, Sven, & Olson, Jonas. (2007). Brentano and the Buck-Passers. *Mind*, 116(463), 511–522.
- Ewing, A. C. (1948). *The definition of good*. Oxford: Routledge.
- Fritzson, F.-A. (2014). *Value grounded on attitudes: Subjectivism in value theory* (doctoral dissertation). Lund: Media-Tryck.
- Gertken, J., & Kieseewetter, B. (2017). The right and the wrong kind of reasons. *Philosophy Compass*, 12(5), e12412.
- Howard, C. (2018). Fittingness. *Philosophy Compass*, 13(11), e12542.
- Jacobson, D. (2011). Fitting attitude theories of value. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2011 Edition). Stanford, CA: The Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University.
- McHugh, Conor, & Way, Jonathan. (2016). Fittingness First. *Ethics*, 126(3), 575–606.
- Olson, J. (2004). Buck-passing and the wrong kind of reasons. *The Philosophical Quarterly*, 54(215), 295–300.
- Orsi, F. (2013). Fitting attitudes and solitary goods. *Mind*, 122(487), 687–698.
- Parfit, D. (2011). *On what matters*. Oxford: Oxford University Press.
- Rabinowicz, W., & Rønnow-Rasmussen, T. (2004). The strike of the demon: On fitting pro-attitudes and value. *Ethics*, 114(3), 391–423.
- Reisner, Andrew E. (2015). Fittingness, value and trans-world attitudes. *Philosophical Quarterly*, 260, 1–22.
- Rønnow-Rasmussen, T. (2011). *Personal value*. Oxford: Oxford University Press.
- Rowland, R. (2014). Dissolving the wrong kind of reason problem. *Philosophical Studies*, 172(6), 1455–1474.
- Rowland, R. (2019). *The normative and the evaluative: The buck-passing account of value*. Oxford: Oxford University Press.
- Scanlon, T. (1998). *What we owe to each other*. Belknap: Harvard.
- Skorupski, J. (2007). Buck-passing about goodness. In T. Rønnow-Rasmussen, et al. (Eds.), *Hommage a' Wlodek: Philosophical papers dedicated to Wlodek Rabinowicz*. Lund: Lund University.
- Väyrynen, P. (2013). Grounding and normative explanation. *Aristotelian Society Supplementary*, 87(1), 155–178.
- Way, J. (2012). Transmission and the wrong kind of reason. *Ethics*, 122(3), 489–515.