

Reasons as Reasons for Preferences*†

R.A. Rowland
University Leeds

Abstract. I argue that all reasons for actions and attitudes consist in reasons for preferences; call this view *RP*. According to *RP*, reasons for \mathcal{A} to believe that p just consist in reasons for \mathcal{A} to prefer their believing that p to their not believing that p , and reasons for \mathcal{A} to have a pro-attitude or perform an action just consist in reasons for \mathcal{A} to prefer that she has that attitude/performs that action. I argue that we have strong reason to accept *RP* because we can explain a correlation between reasons for preferences and other reasons only if we accept *RP*. I argue that no objections undermine *RP* and that *RP* has interesting implications for the reasons for attitudes there are and for reasons fundamentalism.

I

In this paper I argue that

RP. What it is for R to be a reason for \mathcal{A} to ϕ (except where ϕ is a preference or desire) is just for R to be a reason for \mathcal{A} to prefer \mathcal{A} 's ϕ -ing to \mathcal{A} 's not- ϕ -ing (and for \mathcal{A} to be able to ϕ on the basis that R).

According to *RP*, reasons for belief, reasons for action, and reasons for pro- and con-attitudes, such as to admire, to intend, and to condemn, just consist in reasons for preferences. According to *RP*, reasons for \mathcal{A} to believe that p just consist in reasons for \mathcal{A} to prefer their believing that p to their not believing that p , reasons for \mathcal{A} to admire someone just consist in reasons for \mathcal{A} to prefer their admiring them, and reasons for us to perform an action just consist in our reasons to prefer our performing that action.¹

RP needs a few clarifications. First, what is the notion of preference at play here? Preferences are subjective mental states, as are approvals and desires. But unlike all desires, all

* This is a preprint of a paper forthcoming in *American Philosophical Quarterly*. Please cite the published version of this paper when it is available.

† I would like to thank Guy Fletcher, Conor McHugh, Jonathan Way, and three referees for the journal for extremely helpful comments on a previous draft of this paper.

¹ See Piller (2006) for a somewhat similar view.

preferences are comparative.² I assume for the sake of the paper that reasons for desires can be reduced to reasons for preferences. But if you think that desires are more fundamental than preferences, then understand *RP* as an analysis of all reasons in terms of reasons for desires.³

Second, what kind of account or analysis does *RP* provide? *RP* is a reductive property analysis according to which the property of being a reason for an action is reduced to the property of being a reason for a preference; and, similarly, the property of being a reason for a belief (and being a reason for all other attitudes) consists in the property of being a (type of) reason for a preference.⁴

Third, it might seem that *RP* is circular. Since it analyses the property of being a reason to perform an action and have certain attitudes in terms of the property of being a reason to have another kind of attitude. But *RP* is not a reductive account of what it is to be a reason in general. It is just a reductive account of reasons other than reasons for preferences. So it is not circular. Similarly, accounts of moral reasons in terms of prudential reasons are not circular. Such accounts provide an account of one type of reason in terms of another.

Fourth, *RP* is compatible with both reductive and non-reductive views about the nature of the property of being a reason. For instance, it is compatible with *RP* that for there to be a reason for *A* to prefer *X* to *Y* is just for it to be fitting for *A* to prefer *X* to *Y*. But it is also compatible with *RP* that the property of being a reason to have a preference is fundamental and irreducible.⁵

Fifth, *RP* does not entail instrumentalism about practical and epistemic reasons. For *RP* is compatible with the view that there are non-instrumental reasons for preferences, that there are reasons to prefer things for their own sake, and that there are reasons to prefer that we believe that *p* regardless of the further good consequences of our believing that *p* (and so that there are reasons for us to believe that *p* that outstrip the good consequences of our believing that *p*).

² For a more thorough account of the nature of preferences and related relevant issues see Hansson and Grüne-Yanoff (2018).

³ I hope to be agnostic regarding debates about the nature of desire; for an introduction to such debates see Schroeder (2015).

⁴ On what reductive property analyses are see, for instance, Schroeder (2007: ch. 4, esp. 67-72).

⁵ For discussion of such reductive and non-reductive views see McHugh and Way (2016), Kearns and Star (2009), and *infra* note 30.

It might seem that this is not right for reasons for preferences are always reasons to prefer one's obtaining a goal or objective, and if this is right, then *RP* seems to entail instrumentalism about practical and epistemic reasons. However, *RP* does not imply this. Consider attitudes. There can be non-instrumental reason to prefer having an attitude. For instance, as I'll explain in section 2, we plausibly have reasons to prefer that we have beliefs that fit our evidence even if nothing good comes from this, and we have reason to prefer that our admiration is apt, even if nothing good comes from such admiration; these need not be reasons to prefer having the relevant attitude for the reason that having it serves some objective or goal of ours.⁶

Finally, *RP* does not entail that whenever we have a reason to prefer our ϕ -ing to our not- ϕ -ing, then we have a reason to ϕ . The parenthetical caveat in *RP* ensures that *R* is a reason for us to ϕ only if there is a reason for us to prefer that we ϕ and we can ϕ on the basis that *R*. In this paper I want to be non-committal about whether this caveat should be included in *RP*, which is why I include it in parentheses. Some hold that *R* can be a reason for us to ϕ even if we cannot ϕ on the basis that *R*. For instance, some hold that the fact that a demon will punish us if we do not admire it is a reason for us to admire the demon even if we cannot admire it for this reason. And some hold that there are Pascalian reasons to believe in God even if we cannot believe in God for such reasons.⁷ *RP* can be made compatible with this view by removing the parenthetical caveat from it. This paper isn't intended to settle the debate between skeptics about reasons of the wrong kind, such as the reasons to admire the demon and believe in God in the aforementioned cases, and those who believe that there are such reasons (of the wrong kind). And it doesn't need to either, since whether we are wrong kind of reasons skeptics or not we can still accept that all reasons for attitudes and actions are just reasons for preferences.⁸

In sections 2-3 I argue that we have a strong reason to accept *RP* because we can explain a correlation between reasons for preferences and other reasons only if we accept *RP*. In section 4 I argue that no objections undermine *RP*. And in section 5 I argue that *RP* has interesting implications regarding debates about reasons fundamentalism and debates about what reasons for attitudes there are.

⁶ I will say a little more about how *RP* does not imply a form of instrumentalism or pragmatism, in *infra* note 25.

⁷ See, for instance, Howard (2019).

⁸ For discussion of wrong kind of reasons skepticism see Skorupski (2007: 9-12), Way (2012), Rowland (2015), and McHugh and Way (2016: 580-583, 594-595); see also Shah (2006).

II

There seems to be a striking correlation between reasons for us to ϕ and reasons for us to prefer that we ϕ . Namely, it seems that

Striking. At least very often, if there is a reason for us to ϕ , then there is reason for us to prefer that we ϕ rather than that we do not ϕ .

To show that *Striking* holds I'll consider reasons for action, reasons for pro-attitudes, and reasons for belief (and these reasons' relationship to reasons for preferences) in turn.

Reasons for Action

First, consider the relationship between reasons for *action* and reasons for preferences. If there is a reason for you to go to Melbourne rather than Canberra (because of its better music and food scenes), then there is a reason for you to prefer going to Melbourne rather than Canberra (provided by its better music and food scenes). If the fact that I will get more work done if I stay home is a reason for me to stay home today rather than go to the office, then the fact that I will get more work done is a reason for me to prefer to stay at home today rather than go to the office. Similarly, it seems that if there is a reason for us to refrain from pushing the heavy man off of the bridge in a trolley case, then there is reason for us to prefer that we refrain from pushing the heavy man off of the bridge in this case.⁹ So, at least very often, if there is a reason for us to perform action X rather than action Y , then we have reason to prefer that we perform action X rather than action Y .

Reasons for Pro-Attitudes

Similarly, it seems that very often we have reasons to prefer our having the pro-attitudes that we have reasons to have. It seems that we have reasons to prefer that we admire admirable people such as virtuous and creative people. And people who are admirable are people whom we have reasons to admire. So, it seems that we have reasons to prefer that we admire those whom we have reasons to admire. Suppose that we have reasons to admire Derek Parfit. In this case it seems that we have reason to prefer our admiring Parfit to our not admiring Parfit; for by admiring Parfit we

⁹ See Way (2013: 37).

are responding aptly to Parfit's admirability. Similarly, it seems that we have reasons to prefer that we desire things that we have reasons to desire. We have reasons to desire health, happiness, and freedom and we have reasons to prefer that we desire these desirable things. For by desiring these things we are responding appropriately to that which is desirable. Similarly, we have reasons to desire more free and equal societies and we have reasons to prefer our desiring more free and equal societies. To clarify, my thought here is not that we have instrumental reasons to prefer our aptly desiring the desirable and admiring the admirable—though we do also have such instrumental reasons. But rather that we have reasons to prefer our desiring the desirable for its own sake: we have reason to prefer our responding correctly to all our reasons for its own sake.

It seems that to the extent that we have reason to desire something we have reason to prefer that we desire that thing to the same extent. For instance, suppose that we can pursue three careers that are differentially desirable. Suppose that: (a) the least desirable career is a career working in Walmart because it will be poorly paid, unfulfilling, and not pleasurable; (b) the moderately desirable career is a career as a school teacher because it would give us a moderately fulfilling existence, a moderate amount of pay, and a moderate amount of happiness; and (c) the most desirable career is a career as an Ivy League professor, which would be very desirable because it would give us a very fulfilling existence, a large amount of pay, and a great amount of happiness. It seems that we have reasons to prefer our wanting the career as a school teacher over the career with Walmart and we have reasons to prefer our wanting the Ivy League career over a career as a school teacher.

So, it seems that very often, if there is a reason for us to have a pro-attitude towards *X*, then there is a reason for us to prefer that we have this pro-attitude towards *X*. This conclusion also seems to follow from Thomas Hurka's recent claim, which he plausibly ascribes to Aristotle, Brentano, and Ross amongst others, that it is good in itself for us to have pro-attitudes towards that which is of value: it is good for us to love, admire, approve of, and desire the good.¹⁰ For: (a) we have reasons to love, admire, approve of, and desire the good for its own sake; and (b) if something's obtaining would be good, then we have a reason to prefer its obtaining, so if it's good for us to do something then, we have a reason to prefer that we do it.¹¹ And so, in this case, very often we have reasons to prefer that we admire that which, and those whom, we have reason to

¹⁰ Hurka (2001: ch. 1)

¹¹ On (b) see Scanlon (1998: 97), Suikkanen (2009), and Way (2013).

admire for its own sake, love that which, and those whom, we have reasons to love for its own sake, and desire that which we have reason to desire for its own sake.¹²

Reasons for Belief

It seems that we have reasons to prefer that we have an understanding of our surroundings and that we have beliefs that constitute knowledge and that fit with our evidence. And for most of the particular beliefs that we have reasons to have we have reasons to prefer that we have these beliefs. For instance, we have reasons to prefer that we have the true beliefs that we have reasons to have about how to navigate our city, which restaurants and coffee shops cater to our preferences, and how to make those whom we love happy.

We have multiple reasons to prefer that we believe that which we have good reason to believe. Instrumentalists' account of reasons for belief entails that we have reasons to prefer that we believe that which we have reason to believe. For instance, according to one instrumentalist,

forming beliefs on the basis of one's evidence is a good way—perhaps the best way, given our cognitive limitations—of forming true beliefs...And, up to a point, having true beliefs is of great practical utility. This is because having true beliefs is a good means of facilitating the goals that one possesses, or should possess.¹³

So, according to instrumentalists, we have instrumental reasons to prefer that we believe in line with our evidence and so believe that which we have reason to believe.

It is also plausibly preferable in itself for us to have (at least non-trivial) beliefs that provide us with an understanding of our world or surroundings. For instance, consider the lives of Diane and Ellie. Both Diane and Ellie lead happy, autonomous lives in which they achieve a lot and have lots of meaningful relationships. Diane and Ellie's lives, let's suppose, are identical from the inside. However, Diane has far fewer true beliefs, far less understanding, and far less knowledge than Ellie. Fill out this case however you want. For instance, suppose that Diane's entire upbringing

¹² What about in cases of outweighed reasons for pro-attitudes? For instance, suppose that there are some reasons to admire my friend, she's smart and funny, but other aspects of her character outweigh the reasons to admire her. However, in this case there is reason to prefer admiring my friend; admiring her is preferable in two ways because it would involve responding to these reasons to admire her. But these reasons to prefer admiring her are outweighed by other features of her character, so there's not all things considered reason to prefer admiring her. See also Maguire (2017).

¹³ Cowie (2014: 4007). See also Schroeder (2007: ch. 6).

was spent in a Truman-show-esque deception—but a deception involving others being deceived as well so that she did have true friendships—that she never discovered was a deception. It seems that Diane’s lack of understanding is worthy of regret even though this lack of understanding did not harm her in any other way. So, it seems that Ellie’s life is preferable to Diane’s. And it seems that if we had to choose to live Diane’s or Ellie’s life we would choose to live Ellie’s life. Furthermore, most epistemologists hold that either having true beliefs or having understanding and/or wisdom is of epistemic value.¹⁴ But in this case it is preferable for us to have beliefs that are true or (partially) constitute or lead to understanding or wisdom.

It might seem that all this only shows that there are reasons for us to prefer that we believe that p if we have strong reason to believe that p or if p is true (or true given our evidence). However, do we really have reasons to prefer that we believe that p if we do not have sufficient reason to believe that p ? It seems that we do. Right now, October 2020, there is some reason to believe that Biden will win the election and become President of the United States and there is some reason to believe that Biden will win the election but Trump will nonetheless find a way to cling onto power. It does seem that it’s *pro tanto* preferable to believe one of these things on the basis of the relevant evidence, since believing one of these things on the basis of the relevant evidence involves an understanding of the relevant evidence. But, assuming that the evidence is not quite sufficient to justify either believing that Biden will become President after winning the election or that Trump will retain the presidency even after losing the election (or believe that some other outcome will obtain), we have most reason to prefer that we do not believe either of things because in so believing these things we would be putting more weight on certain pieces of evidence than we should.

Consider cases in which there is some reason to believe that p but more reason to believe that not- p . For instance, suppose that the fact that a policy would involve building a new vibrant arts district in your city is a reason to believe that the policy is a good one. But this reason to believe this is outweighed by the fact that the building of this arts district would involve the destruction of people’s homes against their wills and the gentrification of an area that would make many people’s lives more expensive and difficult. In this case, it is in some ways preferable to believe that the policy would be a good one because it does have these indisputably good features but there is more reason to prefer that one does not believe this because of the worse features of the policy that outweigh the good features.

¹⁴ See, for instance, Dunn (2015, section 2), and the references therein.

So, when there is some (outweighed or insufficient) reason to believe that p there is still a reason to prefer that one believe that p . This view also fits well with the idea that understanding is of epistemic value.¹⁵ Since, if one believes that p on the basis of some evidence, but this evidence is insufficient to establish that p or is outweighed by other evidence, then one's belief that p still involves a degree of understanding, and to this extent is a belief that it is preferable to have. But the degree of understanding that one's belief that p involves is not as great as the degree of understanding that refraining from believing that p or believing that not- p would involve. So, although there is some reason to prefer that one believes that p (provided by the degree of understanding this would involve) there is stronger reason to prefer that one does not believe that p or believes that not- p (due to the greater understanding that believing that not- p or not believing that p would involve).¹⁶

III

I've been arguing that

Striking. At least very often, if there is a reason for us to ϕ , then there is reason for us to prefer that we ϕ rather than that we do not ϕ .

RP explains why *Striking* holds. For according to *RP*, reasons for us to ϕ just consist in reasons for us to prefer that we ϕ rather than that we do not ϕ .

I'll consider three other ways of explaining *Striking*. The first way would be to reduce reasons for preferences to reasons for other actions and attitudes. This view would explain *Striking* in the opposite way to the way in which *RP* explains *Striking*: *RP* explains *Striking* by reducing its antecedent (reasons for us to ϕ) to its consequent (reasons for us to prefer that we ϕ); this view would explain *Striking* by reducing its consequent to its antecedent.

¹⁵ *Supra* note 14

¹⁶ It has been put to me that this strategy is less plausible in the case of reasons to believe that $2+2=4$. But this is not the case: if I have epistemic reason to believe that $2+2=4$, then I have reason to prefer that I believe that $2+2=4$ rather than that I do not believe this.

However, we cannot analyse all reasons for preferences in terms of reasons for action and reasons for other attitudes. There is a reason for us to prefer that Trump does not start World War III to his not doing so, but this can't plausibly be reduced to a reason for action that we have. There is a reason for us to prefer that we turn ourselves into a killer whale for one minute if a demon will punish us if we do not but since we cannot turn ourselves into a killer whale for one minute there is no reason for us to do so.¹⁷ So, the reason to prefer that we turn ourselves into a whale cannot be understood in terms of a reason for an action. And there is a reason for us to prefer that we believe that there are square circles if a demon will punish us if we do not but if we cannot believe this, there is no reason for us to believe this. So, if we cannot believe this, there is no way of analyzing this reason to prefer in terms of a reason for belief. So, we cannot plausibly explain *Striking* by reducing its consequent (its right-hand side) to its antecedent (its left-hand side).

Alternatively, it might seem that a first-order strategy can be adopted to explain *Striking*. Adopting this strategy would involve giving a first-order account of our reasons for preferences, reasons for belief, and reasons for other actions and attitudes and holding that it is the overlap that results from the combination of these accounts that leads to *Striking*. For instance, it might be held that happiness, achievement, and understanding give rise to our reasons for preferences, reasons for belief, and reasons for other actions and attitudes. According to this view, we have reasons to believe things in order to gain understanding and we have reasons to prefer that we believe things because we have reasons to prefer that we gain this understanding. And, similarly, we have reasons to perform actions that promote happiness and our own achievement as well as reasons to prefer that we perform actions that promote happiness and our own achievement. So, it is the relationship between reasons for preferences and happiness, achievement and understanding, and the relationship between reasons for belief and action and happiness, achievement, and understanding that explains *Striking*.

However, this first-order strategy cannot adequately explain *Striking*. It seems that whatever first-order view about our reasons for actions and attitudes that one adopts it should entail that we have reasons to prefer that we have the attitudes that we have reason to have and perform the actions that we have reason to perform. For instance, whatever first-order view one holds about the footbridge trolley case one should hold that our reasons for action and our reasons for preferences lineup: if you hold the view that we have most reason to push the heavy man to his death to save five lives, you should hold that we have most reason to prefer that we push the

¹⁷ See, for instance, Streumer (2007).

heavy man; if you hold the view that we have most reason to prefer that we refrain from pushing the heavy man, you should hold the view that we have most reason to refrain from pushing the heavy man. *Striking* points to a connection between reasons to ϕ and reasons to prefer that we ϕ that holds regardless of which first-order view that we hold. So, *Striking* cannot be explained by a first-order view of our reasons to prefer and our reasons to have other attitudes and perform other actions; *Striking* points to a plausible constraint on such first-order views.

A referee has suggested a third way of explaining *Striking*, according to which, we explain *Striking* by holding:

(A) Whenever there is a reason for us to ϕ (where ϕ is an action or attitude other than a desire or preference) and also a reason for us to prefer that we ϕ , the latter reason is explained by the former.

On this view, whenever there is a reason to perform act A and also a reason for us to prefer that we perform act A, the reason to prefer is explained by the reason to act and whenever there is a reason for us to admire and a reason for us to prefer that we admire, the reason for us to prefer that we admire is explained by the reason to admire.

However, (A) does not explain *Striking* for it only claims that the reasons to prefer that we ϕ that are correlated with reasons to ϕ are explained by the latter reasons. But it does not tell us why they are explained by the latter reasons: in order for this kind of view to explain *Striking* we need an explanation of why it is the case that very often reasons to ϕ give rise to reasons to prefer that we ϕ .¹⁸

One idea that might be easily combined with (A) to explain *Striking* would be that

¹⁸ Similarly, suppose that there is a striking correlation between my beliefs about an island and the facts about this island. One gives no explanation of this correlation if one claims that my beliefs are explained by the facts about the island but one cannot explain how this is the case (e.g. if one knows that I've not been there, nor have my friends, I've not watched a TV show or read a book about the island, etc); see Rowland (2019, pp. 56-59)].

Analogously, Leary (2017, pp. 86-87) argues that attempts to explain the supervenience of the normative on the non-normative by invoking particular moral principles that state necessary correlations between the normative and the non-normative just restate more specific supervenience claims rather than explaining supervenience.

(B) We very often have *pro tanto* reason to prefer that we respond to, and act in line with, our reasons.

The combination of this claim and (A) would seem to explain *Striking*. For on this view what explains *Striking* is that reasons to ϕ give rise to reasons to prefer that we ϕ because—regardless of what reasons we have—we have reasons to prefer that we respond to our reasons and act in line with what we have reason to do.

However, *RP* provides an explanation of *Striking* that is simpler than this explanation, for *RP* holds that there is only one fundamental kind of reason, reasons for preferences—we will discuss this virtue of *RP* in section 5—this explanation does not yield such simplicity. Furthermore, (B) itself states a striking correlation that requires explanation. So, in order for the combination of (A) and (B) to explain *Striking*, (B) still needs to be explained.

One natural explanation of (B) would be that

(C) We have reason to prefer that we are rational and responsive to reasons.

But we cannot use the combination of (A), (B), and (C) to explain *Striking*. To see this suppose that objectivism about reasons is correct, that is, that whether we have reasons to perform an action depends on the objective facts even if these are outside our epistemic perspective.¹⁹ Facts O1, O2, and O3 give you reason to perform actions 1-3, but you are blamelessly unaware of O1-3. Given your lack of awareness, your refraining from performing these actions would manifest no rational failing or lack of responsiveness to reasons. In which case, (C) cannot explain why there is reason for us to prefer that we perform actions 1-3. But if we are objectivists who hold that there are objective reasons to perform 1-3 in virtue of O1-3, we should also hold that there are reasons for us to prefer that we perform 1-3; this is part of the phenomenon of covariation that *Striking* denotes, and *RP* explains. For instance, suppose that your flipping your light switch will, by an unforeseen set of events, trip a switch that will catapult an innocent bystander in front of a trolley track bringing it to a stop, killing them, and saving five. Deontologists who are objectivists (rather than perspectivalists) about reasons hold (and should hold, see section 1 above) that you have most reason to refrain from flipping the light switch and most reason to prefer that you refrain from flipping the light switch. But your not refraining from flipping the light switch

¹⁹ For discussion see, for instance, Lord (2015).

would not demonstrate a lack of rationality or responsiveness to reasons. So, we cannot explain (B) by invoking (C). In which case we cannot fully explain *Striking* by appeal to (A) and (B) unless we can otherwise explain (B) without accepting *RP*; but it is hard to see what else would explain (B) other than (C) or *RP*.²⁰

So, there are some issues with explaining *Striking* by invoking (A). I don't hope to have shown that it is impossible to explain *Striking* in this way. But I do hope to have shown that it is not clear how to provide such an explanation of *Striking* that is both plausible and a genuine explanation of *Striking*. Given that no such explanations have in fact been given in the literature as yet, for the time being, it is reasonable to hold that *RP* explains *Striking* better than any alternative view.²¹ But *RP* is a striking correlation and striking correlations call out for explanation. We want to know—or should want to know—how come reasons to ϕ always or very often come together with reasons to prefer that we ϕ . (Similarly, if people from a particular town were always, or very often, good at a particular sport, we would want to know what explains this; normally we'd be able to find the explanation quite quickly (e.g. a particularly good trainer or training facility). And, as I've been arguing we can fully explain *Striking* only if we accept *RP*. Furthermore, if we can fully explain a particular striking correlation by accepting view V but we cannot fully explain this particular striking correlation if we do not accept V , other things equal, we should accept V .²² So, other things equal, we should accept *RP*.

IV

²⁰ Perhaps the main motivation for trying to explain *Striking* by invoking (A) rather than *RP* is that it seems that our reasons to perform actions are explanatorily prior to our reasons to prefer that we perform those actions; see Dancy (2018, pp. 89-90). But it is compatible with *RP* that reasons for action in some sense explain our reasons for preferences. Since *RP* identifies our reasons for action with a subset of reasons for preferences; it is also, of course, consistent with *RP* that some things are worth doing, and this is what—and/or the facts that make these things worth doing are what—makes it the case that we have reason to prefer our doing them. *RP* just holds that, metaphysically, what it is to be a reason for an action is to be a particular kind of reason for preference.

²¹ Perhaps a value-first account of what it is to be a reason to have any attitude or perform an action could explain *Striking*. However, Way (2013) and Rowland (2019, ch. 2, 3, 5) have argued that we should reject such a view. Furthermore, even if I can only conclude that we can explain *Striking* only if we accept *RP* or the extremely controversial value-first account of all reasons, this would not diminish the strength of the reason to accept *RP* because it explains *Striking*.

²² See Enoch (2011: 158) and McHugh and Way (2016: 582).

I can imagine three particularly pressing objections to *RP*. First, it might seem that *RP* provides an unattractive view of what it is to be an epistemic reason for belief because our epistemic reasons for belief are just a function of our evidence.²³ However, *RP* is consistent with the view that our reasons for belief are a function of our evidence. In section 2 I argued that we have reasons to prefer that we form beliefs in line with our evidence for its own sake. If this is right, then to the extent that *E* is evidence that *p* (that we have access to) we have reason to prefer that we form our beliefs about *p* in line with *E*.

Furthermore, some have denied that our epistemic reasons are a function of our evidence. And it is a distinctive virtue of *RP* that it accommodates and makes particularly good sense of this debate about whether epistemic reasons are a function of our evidence. Some have claimed that it does not follow from the fact that *E* is evidence that *p* (that we have access to) that *E* is thereby a reason for us to believe that *p*. Those who hold that our evidence does not determine our reasons for belief in this way have claimed that it is extremely intuitive to them that when *p* is a proposition that is entirely trivial or entirely unimportant to us (such as the proposition that there are 339,000 blades of grass in a particular unimportant garden), then we have no reason to believe that *p* even when we have strong evidence that *p*.²⁴ If we accept *RP*, we can explain why it seems intuitive to many, and why many have claimed that, we sometimes have no reason to believe that *p* even when we have strong evidence that *p*. For according to *RP*, our reasons to believe that *p* just consist in our reasons to prefer that we believe that *p* rather than not. And it seems to some that we have no reasons to prefer that we believe entirely trivial propositions even when we have strong evidence for them, since our believing such propositions is of no great epistemic importance; but it seems to others that we have reasons to prefer that we believe entirely trivial true propositions because it is (epistemically) preferable for us to have true beliefs or beliefs that we have evidence for even if these beliefs are about unimportant things.²⁵

²³ See Portmore (2011: 133-135).

²⁴ See Cowie (2014: 4014-4015) and Papineau (2013).

²⁵ A referee has put to me that on this picture, pragmatism, the view that reasons for belief are actually reasons for taking certain actions, looms. But this is not the case. According to *RP*, both reasons for belief and reasons for action are analysed in terms of reasons for preferences. And on this view we might have epistemic reasons to believe that *p* (because we have reason to prefer that we believe that *p*) even if we have no reason at all to act to bring it about that we believe that *p*. Suppose that if *E* is evidence that *p* that *A* is aware of, then that is sufficient to establish that *E* is a reason for *A* to believe that *p*. But suppose that, for instance, we know that all the evidence favours the thesis that we are of average attractiveness, but because of our self-hatred we just cannot form this rational belief; we can only believe that we are very unattractive. In this case, in addition to the reason to prefer that we believe that we are of average attractiveness, because that will make us feel better, we also have reason to prefer that we believe that we are of average attractiveness because having this belief will involve our correctly responding to the evidence. The epistemic reason to believe that we are of average attractiveness is constituted by the latter reason to prefer that we

A second objection concerns wrong kind of reason cases. If a demon will punish us if we do not admire it, there is (a) a reason for us to *prefer that we admire the demon*. According to some, in this case there is also (b) a reason for us to *admire the demon*. Now suppose that the demon threatens Alex. Alex prefers her admiring the demon to her not doing so. But she does not admire the demon. In this case, Alex has conformed with her reason to prefer to admire the demon, but she has not conformed with her reason to admire it: she has conformed with (a) but not (b)—supposing there is such a reason as (b). But *RP* seems to make this impossible because, according to *RP*, if there is a reason for Alex to admire the demon, then this reason just consists in a reason for her to prefer her admiring the demon. And in this case although she seems to have conformed with all her reasons to prefer that she admires the demon she has not conformed with all her reasons to admire the demon.

However, it is not the case that *RP* entails that we comply with our reasons to admire whenever we prefer that we admire. For we should hold that we comply with our reasons to admire *X* only if we admire *X* and we comply with our reasons to prefer that we admire *X* only if we prefer to admire *X*. It is not ad-hoc to make this move for the view that I'm arguing for is not an error theory about reasons to have attitudes and perform actions other than preferences. Rather it is a reductive account of reasons for actions and other attitudes in terms of reasons for preferences. But in this case we want to preserve the way in which reasons for other attitudes and actions are distinct from standard reasons for preferences whilst still maintaining a reductive account of these reasons in terms of reasons for preferences. One way in which reasons to admire (for instance) are distinct from reasons to prefer admiring is that reasons to admire may be subject to a response condition that reasons to prefer to admire are not subject to: *R* may be a reason to admire *X* only if we can admire *X* on the basis of *R*; *R* can be a reason to prefer to admire *X* even if we cannot admire *X* on the basis of *R* (see section 1 above). Another way in which reasons to admire (for instance) are distinct from reasons to prefer admiring is that they have different compliance conditions: if we have most reason to admire *X*, we comply with our reasons to admire only if we admire *X*; and if we have most reason to prefer to admire *X* we comply with our reasons to prefer that we admire *X* only if we prefer to admire *X*. But nevertheless, we can reductively analyse

are of average attractiveness. But we have no reason to bring it about that we believe this, since we cannot bring this about (and we have a reason to perform an action only if we can perform it; see Streumer 2007).

reasons to admire in terms of reasons for preferences and these response and compliance conditions.²⁶

A related objection concerning wrong kinds of reasons has been put to me. It has been put to me that *RP* begs the question against skeptics about wrong kinds of reasons who hold that a reason to prefer to ϕ is not a reason to ϕ . However, wrong kind of reasons skepticism is not centrally the claim that reasons for action do not consist in reasons for preferences. Rather wrong kind of reasons skepticism is the claim that the instrumental benefits of having certain attitudes does not provide us with reasons to have those attitudes. And *RP* is compatible with this claim. Remember that according to

RP. What it is for R to be a reason for A to ϕ (except where ϕ is a preference or desire) is just for R to be a reason for A to prefer A 's ϕ -ing to A 's not- ϕ -ing (and for A to be able to ϕ on the basis that R).

Wrong kind of reasons skeptics hold that there are no reasons to admire a demon who will punish you if you do not admire it because we cannot do so on the basis of this reason; it is impossible to admire a demon on the basis of non-admirability-related considerations.²⁷ But although there is a reason for us to prefer that we admire the demon, wrong kind of reasons skeptics can affirm the parenthetical caveat in *RP* and hold that this does not make it the case that there is a reason to admire it since we cannot admire it on the basis that it will punish us if we do not. So, *RP* seems to be compatible with wrong kind of reason skepticism.

A referee has put a third objection to me. Sometimes A has most reason to ϕ but does not have most reason to prefer that they ϕ . Consider, for example, a parent who has two children Fifi and Ursula. Fifi's jokes are funnier than Ursula's. The parent has most reason to find Fifi's jokes funnier than Ursula's, since the jokes are funnier. But they have most reason to prefer their finding both jokes equally funny because if they (persistently) found Fifi's jokes funnier than Ursula's, this would involve them treating, or lead them to treat their children unequally: it would lead to them

²⁶ Some objections have been made to the view that we should reduce the property of being a reason for action to the property of being a reason for a preference; see Scanlon (1998: 50-55, 88-89). But Portmore (2011: 132, 136-140) shows that these objections fail.

²⁷ See *supra* note 8.

favouring Fifi over Ursula. It seems that *RP* cannot make room for cases like this because it holds that reasons to find funny are identical to our reasons to prefer that we find things funny.

This is an interesting and important line of objection that I can only briefly respond to here. I believe that this case does not present a problem for *RP* for similar reasons to the reasons why wrong kind of reasons skepticism does not present a problem for *RP*. According to wrong kind of reasons skeptics, the only reasons to admire are admirability reasons because they are the only reasons that we can admire things on the basis of; so there is no reason to admire the evil demon. But there is reason to prefer that we admire the demon because more than admirability based considerations give us reason to prefer that we admire it, since we can prefer that we admire the demon based on other considerations such as the consequences of our admiring them. A similar form of explanation explains why in the Fifi/Ursula case some think that there is most reason for the parent to find Fifi's jokes funnier but not most reason for them to prefer finding Fifi's jokes funnier. Namely, only funniness reasons are reasons to find the joke funny, since those are the only kinds of reasons on the basis of which we can in fact find the joke funny. But more than funniness based considerations bear on whether we have most reason to prefer that we find the joke funny because we can prefer/not prefer that we find the joke funny on the basis of a more diverse variety of considerations such as the inequalitarian consequences of our finding the joke funny. So, it is consistent with *RP* that the weight of the reasons to ϕ is determined by the set of reasons to prefer that we ϕ , *set A*, that we can ϕ on the basis of, but that the set of reasons to prefer that we ϕ can be a greater set of reasons than *set A*. In this case, it is consistent with *RP* that in some cases we can have most reason to ϕ even though we do not have most reason to prefer that we ϕ .

V

I've argued that there is a significant theoretical advantage to accepting *RP* and that *RP* does not face any serious problems. So, we should accept *RP*. There are at least three interesting implications of *RP*.

RP provides a unified explanation of why we have reasons to have and do a diverse variety of things: perform a range of actions, have beliefs, have desires, approve of and admire things, be averse to, and fear things. According to *RP*, we have reasons to do such a diverse variety of things

because we have reasons to prefer our doing this diverse variety of things. This feature of *RP* provides us with a reason to accept it. Just as the fact that rule-consequentialism and contractualism provide a unified explanatory account of why we have moral duties to perform a diverse range of actions (e.g. give to charity, keep our word, tell the truth, avoid harming others) gives us (*pro tanto*) reason to accept these views.²⁸

The fact that *RP* provides a unified explanation of why we have reasons to have a variety of attitudes and do a diverse variety of things has important implications for reasons fundamentalism. According to reasons fundamentalism, the property of being a reason is an irreducible property that cannot be analysed in other terms and other normative properties such as value, ought, and fittingness can be analysed in terms of the property of being a reason.²⁹ Many have claimed that if we accept reasons fundamentalism, then we cannot provide an illuminating account of reasons. For instance, Kearns and Star claim that we can explain why reasons for beliefs and actions are the same kind of thing only if we accept their view that all reasons for ϕ -ing are evidence that we ought to ϕ ; we cannot explain this, they argue, if we accept competing views such as the view that the property of being a reason is a primitive irreducible property.³⁰ But if we accept *RP*, then we can explain why reasons for beliefs and reasons for actions are the same kind of thing, for according to *RP* these reasons just consist in reasons for preferences. And more generally, as I've been explaining, we can give an illuminating account of the nature of all reasons, and why we have reasons to do the diverse variety of things and have the diverse variety of attitudes that we have if we accept *RP*.

We might be able to provide an account of why we have reasons to do the diverse variety of things that we have reasons to do without accepting *RP* if we are not reasons fundamentalists; if we analyse reasons in terms of fittingness, for instance, we can hold that we have reasons to do the diverse variety of things that we have reasons to do because it is fitting to do these things. But if we accept reasons fundamentalism, it is unclear how we could explain why we have reasons to

²⁸ See Shafer-Landau (2015: 17) and Hooker (2016: section 7). To clarify, I don't mean to claim that rule consequentialism or contractualism are reductive accounts of what it is for something to be a moral duty; for discussion see Scanlon (2007). These views provide a unified explanation of our duties to perform a diverse range of actions by providing a further explanatory principle that is deeper than the other moral principles; *RP* provides a unified explanation by holding that there are reasons for us to prefer that we have the attitudes and perform the actions that we have reasons to have and perform and that these reasons for preferences are more fundamental than the other reasons.

²⁹ See Skorupski (2010), Parfit (2011), and Scanlon (2014).

³⁰ Kearns and Star (2009: 219-222). Others who have claimed that reasons fundamentalism is unilluminating include Väyrynen (2011) and Wedgwood (2015).

do the diverse variety of things that we have reasons to do without accepting *RP*. So, reasons fundamentalists can explain why we have reasons to do the diverse variety of things we have reason to do only if they accept *RP*. So reasons fundamentalists have a good reason to accept *RP*. And, given *RP*, it is not the case that if we accept reasons fundamentalism, we must accept an unilluminating account of what reasons are.

Second, *RP* gives us good reason to reject the view, which some, such as Barry Maguire, have proposed that although there are reasons for action and belief there are no reasons for (affective) pro-attitudes such as reasons to want.³¹ If *RP* holds, then reasons for action and reasons for belief just are reasons for pro-attitudes, so this view could not be correct.³² *RP* does not, however, entail that no first-order arguments can be made against the existence of particular subsets of reasons for pro-attitudes or other actions and attitudes. It is consistent with *RP*, for instance, that we should hold that there are no pragmatic reasons for belief or admiration and that arguments for these claims can be made on the basis that, for instance, (as I discussed in section 1) we cannot admire or believe things on instrumental grounds.

Relatedly, a third interesting implication of *RP* is that it provides us with a new way of getting a grasp on particular first-order arguments regarding reasons to have particular attitudes. Gilbert Harman argues that we have no reasons to feel guilty because it is not desirable to feel guilty.³³ And others have argued that we have no reasons for other reactive attitudes such as blame and resentment on the basis that these emotions are pernicious in various ways.³⁴ *RP* allows us to make sense of how the undesirability or perniciousness of having particular attitudes can bear on whether there are (right kind of) reasons for those attitudes. For according to *RP* whether there are reasons for us to have particular attitudes is a function of whether it is preferable (broadly construed) for us to have these attitudes.

University of Leeds

References

³¹ See Maguire (2018).

³² Faraci (2020) undermines Maguire's arguments for his view.

³³ See Harman (2009)

³⁴ See Sher (2006: viii) and Nussbaum (2016)

- Cowie, Chris (2014). 'In Defence of Instrumentalism About Epistemic Normativity'. *Synthese* 191: 4003-4017.
- Dancy, Jonathan (2018). *Practical Shape*. (Oxford, UK: Oxford University Press).
- Dunn, Jeffrey (2015). 'Epistemic Consequentialism'. *Internet Encyclopedia of Philosophy*. <http://www.iep.utm.edu/>
- Enoch, David (2011). *Taking Morality Seriously: A Defense of Robust Realism*. (Oxford, UK: Oxford University Press).
- Faraci, David (2020). 'We Have No Reason to Think There Are No Reasons for Affective Attitudes'. *Mind* vol. 128, no. 513: pp. 225-234.
- Hansson, Sven Ove and Grüne-Yanoff, Till (2018). 'Preferences', *Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2018/entries/preferences/>>.
- Hooker, Brad (2016). 'Rule Consequentialism'. Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL = <https://plato.stanford.edu/archives/win2016/entries/consequentialism-rule/>
- Howard Chris (2019). 'The Fundamentality of Fit'. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics Volume 14*. (Oxford, UK: Oxford University Press).
- Hurka, Thomas (2001). *Virtue, Vice, and Value*. (Oxford, UK: Oxford University Press).
- Kearns, Stephen, and Daniel Star (2009). 'Reasons as Evidence'. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics Volume 4*. (Oxford, UK: Oxford University Press).
- Leary, Stephanie (2017). 'Non-naturalism and Normative Necessities'. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics Volume 12*. (Oxford, UK: Oxford University Press).
- Lord, Errol (2015). 'Acting for the Right Reasons, Abilities, and Obligation'. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics Volume 10*. (Oxford, UK: Oxford University Press).
- Maguire, Barry (2016). 'The Value-Based Theory of Reasons'. *Ergo* vol 3: pp. 233-262.
- (2018). 'There Are No Reasons for Affective Attitudes'. *Mind* vol. 127: pp. 779-805.
- McHugh, Conor and Jonathan Way (2016). 'Fittingness First'. *Ethics* vol. 126: pp. 575-606.
- Nussbaum, Martha (2016). *Anger and Forgiveness*. (Oxford, UK: Oxford University Press).
- Papineau, David (2013). *There are no norms of belief*. In Timothy Chan (sd.), *The aim of belief*. (Oxford, UK: Oxford University Press).
- Parfit, Derek (2011). *On What Matters Volume 1*. (Oxford, UK: Oxford University Press).
- Piller, Christian (2006). 'Content-Related and Attitude-Related Reasons for Preference'. *Royal Institute of Philosophy Supplement* vol. 59: pp. 155-182.
- Portmore, Douglas (2011) 'The Teleological Conception of Practical Reasons'. *Mind* vol. 120: pp. 117-153.
- Reisner, Andrew (2015). 'Fittingness, Value, and Trans-World Attitudes'. *Philosophical Quarterly* vol. 65, no. 260: pp. 464-485.
- Rowland, Richard (2015). 'Dissolving the Wrong Kind of Reason Problem'. *Philosophical Studies* vol. 172, no. 6: pp. 1455-1474.
- (2019). *The Normative and the Evaluative: The Buck-Passing Account of Value*. (Oxford, UK: Oxford University Press).
- Scanlon, T.M. (1998). *What We Owe To Each Other*. (Cambridge, MA: Harvard University Press).
- (2007). 'Wrongness and Reasons: A Re-examination'. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics Volume 2*. (Oxford, UK: Oxford University Press).
- (2014). *Being Realistic about Reasons*. (Oxford, UK: Oxford University Press).
- Schroeder, Mark (2007). *Slaves of the Passions*. (Oxford, UK: Oxford University Press).
- Schroeder, Tim (2015). 'Desire'. In Ed Zalta (ed.), *Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/entries/desire/#Bib>
- Shah, Nishi (2006). 'A New Argument for Evidentialism'. *Philosophical Quarterly* vol. 56, no. 225: pp. 481-498.

- Sher, George (2006). *In Praise of Blame*. (Oxford, UK: Oxford University Press).
- Shafer-Landau, Russ (2015). *The Fundamentals of Ethics* (Third Edition). (Oxford, UK: Oxford University Press).
- Skorupski, John (2007). 'Buck-Passing About Goodness.' In D. Egonsson, J. Josefsson, B. Petersson, & T. Rønnow-Rasmussen (eds.), *Hommage a Wlodek: Philosophical Essays in Honour of Wlodek Rabinowicz*. Lund University.
<http://www.fil.lu.se/hommageawlodek/site/papper/SkorupskiJohn.pdf>
- (2010). *The Domain of Reasons*. (Oxford, UK: Oxford University Press).
- Streumer, Bart (2007). 'Reasons and Impossibility'. *Philosophical Studies* vol. 136, no. 3: pp. 351-384.
- Suikkanen, Jussi (2009). 'Buck-Passing Accounts of Value'. *Philosophy Compass* vol. 4: pp. 768-779.
- Väyrynen, Pekka (2011). 'A Wrong Turn to Reasons?' In Michael Brady (ed.), *New Waves in Metaethics*. (Basingstoke, UK: Palgrave).
- Way, Jonathan (2012). 'Transmission and the Wrong Kind of Reason' *Ethics* vol. 122, no. 3: pp. 489-515.
- (2013). 'Value and Reasons to Favour'. In Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics Volume 8*. (Oxford, UK: Oxford University Press).
- Wedgwood, Ralph (2015). 'The Pitfalls of "Reasons"'. *Philosophical Issues* vol. 25: pp. 123-143.