

Runaway reasons

Fabienne Peter, f.peter@warwick.ac.uk

September 2024

Forthcoming, *Oxford Studies in Meta-ethics*

Please cite the published version.

1. Introduction

Reasons are central to normativity. We rely on normative reasons when we deliberate about what to do or believe and when we try to justify our actions and beliefs to ourselves and others. But just how central are reasons to normativity? My aim in this paper is to challenge a very common claim about the nature of normativity, which is that reasons are the mark of the normative.¹ As Joseph Raz describes the claim, “[t]he normativity of all that is normative consists in the way it is, or provides, or is otherwise related to reasons” (1999: 67).

The claim that reasons are the mark of the normative is compatible with, but does not presuppose the claim that reasons are explanatorily first. The latter is a claim about the structure of the normative domain – about how the normative realm is constituted, and about the relation between different normative properties and relations. The former is a claim about the nature of normativity. As I understand it here, the normative claim is that the normative realm, however it is constituted, and whatever the relation is between different normative properties or relations, is co-extensive with the realm of reasons. There is no tension between the normativity of reasons and the normativity of other normative properties or relations.

¹ I borrow the expression from Wodak (2020); see also Lord (2023).

Philosophers who defend a reasons-first view of normativity endorse both the explanatory and the normative claims. They hold that other normative properties or relations can be explained in terms of normative reasons, but normative reasons can't be explained in terms of them.² And they endorse the claim that reasons are the mark of the normative because reasons are explanatorily first. If there are no normative properties or relations that are explanatorily prior to reasons, then the normativity of everything normative either consists in reasons or it is otherwise related to reasons.

The so-called Wrong Kind of Reasons problem (Rabinowicz and Rasmussen 2004) puts pressure on the claim that reasons are explanatorily first. Originally focused on the connection between reasons and value, this problem highlights that there appear to be reasons – wrong kind of reasons (WKR) – that fail to appropriately track value. In its more general form, the WKR problem for reasons-first views highlights that reason-based explanations can fail to appropriately track other normative properties or relations. The problem that this creates for the defenders of a reasons-first view is that they must show that they can account for the distinction between WRK and right kind of reasons (RKR) – reasons that track value or other normative properties or relations – in a non-circular way, i.e. without already referring to those properties or relations.

Because the two claims are independent, it's possible to endorse the claim that reasons are the mark of the normative even if the claim that reasons are explanatorily fundamental is rejected. In response to the WKR problem, some philosophers propose a structure of the normative domain in which fittingness is explanatorily prior to reasons (e.g. Chappell 2012; McHugh and Way 2016, 2022;

² Other normative properties or relations typically considered include values, fittingness, virtues, and oughts. Reasons-first views of normativity are defended by Parfit (2011) and Scanlon (2014), among others.

Howard 2019). Normative reasons, they argue, are fittingness-based and fittingness-based reasons are Right Kind of Reasons (RKR): they are normative reasons that appropriately track other normative properties or relations. While they thus reject the explanatory claim, defenders of a fittingness-first view of reasons tend to uphold a version of the normative claim. They hold that normative reasons, as long as they are fittingness-based, or RKR, are the mark of the normative.

My main aim in this paper is to raise a problem for this version of the claim that reasons are the mark of the normative. The problem is not, however, that fittingness-based reasons may not be robustly normative, as some RKR-sceptics have argued.³ This problem obtains even if the normativity of fittingness can be vindicated, as I believe it can. I argue that there are reasons that are anchored in fittingness, and robustly normative as a result, yet fail to be RKR, in the sense that they fail to track the fitting response in a given situation. I call such reasons runaway reasons.

The possibility of runaway reasons implies that there is a normatively significant distinction among normative reasons. There are RKR, which track the fitting response, and runaway reasons. Runaway reasons create a problem for the view that normative reasons are the mark of the normative. The runaway reasons problem is that there are normative reasons that, although fittingness-based, are not RKR. Such reasons have the potential to lead us normatively astray, by failing to track the fitting response in a given situation. If runaway reasons are possible, then the normativity of reasons may be in tension with the normativity of fittingness.

³ See Howard (2019) for a critical discussion of RKR-scepticism. Lord and Sylvan (2019) identify a RKR problem, which is that not all RKR are robustly normative. This problem hinges on doubts about the normativity of fittingness. I address this issue in section 5. The problem I identify here is in some sense the reverse of theirs. It is that not all robustly normative reasons are RKR; see section 7.

The paper is organised as follows. I start with some clarificatory remarks on fittingness and reasons, and on the question of what comes first (section 2). I then move on to consider fittingness-based accounts of normative reasons. In section 3, I introduce Conor McHugh's and Jonathan Way's reasoning account of the relationship between fittingness and reasons (McHugh and Way 2022) and identify a problem with it. In section 4, I introduce and defend an alternative linking principle that explains the relationship between fittingness and reasons in a way that avoids this problem. In section 5, I pause to discuss an objection to all fittingness-based accounts of normative reasons, which is that fittingness is not robustly normative. In section 6, I return to the linking principle to show that some normative reasons are what I call runaway reasons. In section 7, I explain the problem that the possibility of runaway reasons creates for the claim that normative reasons, at least as long as they are fittingness-based, are the mark of the normative. Section 8 concludes.

2. Fittingness and Reasons

I haven't yet said much about what reasons and fittingness are. It's time to rectify that. Regarding reasons, my focus in this paper will be on normative reasons, not on explanatory or motivating reasons. Normative reasons are considerations that play a role in the justification of a response and in the determination of how we ought to respond to a given situation. I follow a standard account here, according to which a normative reason is a consideration that favours some response (Scanlon 1998). The reasons relation, on this understanding, is a relation $R(p, a, x, c)$ between a consideration (or fact) p and some response a for an agent x in circumstances c . For example, that the plant on my desk needs water is a reason for me to water it.

Fittingness is the relation in which an agent's response stands to an object or situation, when the object or situation is worthy of or merits this response.⁴ Relevant responses include actions as well as attitudes. While this understanding of fittingness differs in some ways from some contemporary interpretations, it's in line with more traditional interpretations. According to C. D. Broad (1930 / 2014: 219), for example, "[fi]ttingness or unfittingness is a direct ethical relation between an action or emotion and the total course of events in which it takes place". The fittingness relation, on this understanding, is a relation $F(a, x, c)$ between a response a , an agent x , and some circumstances c – an object or a situation as a whole. In the case of someone admiring a work of art, everything else equal, the relevant circumstances refer to the work of art, the response is admiration, and fittingness is the normative relation between their response and the work of art when the work of art merits admiration.

What we have reason to do (in the broad sense, to include doings and feelings) and what is the fitting response in a given situation might well coincide.⁵ For example, if a child has fallen into the pond right in front of you and is at risk of drowning, and you're the only bystander who can help, that the child is drowning is a reason for you to rescue the child and, let's assume, rescuing the child is also the fitting response in this situation. But the possibility of convergence in the responses that

⁴ This characterisation of fittingness is loosely based on Howard's "gloss" of fittingness (2018). But it's broader in two ways. It includes actions as well as attitudes among the relevant responses, and it considers responses to situations as a whole, not just to isolated objects.

⁵ On a common understanding, which I share, both normative reasons and fittingness apply to beliefs. But to simplify the argument, I will only focus on normative support for actions and practical attitudes in this paper, and bracket beliefs.

they support shouldn't mislead us into thinking that there is no difference between normative reasons and fittingness.

As I see it, there is the following main difference between reasons and fittingness. Normative reasons are contributory and can conflict, whereas fittingness is an overall notion. If an action or attitude is fitting, there is no conflict. If rescuing the child is the fitting action in the pond example I just described, that's the action that fittingness supports in the circumstances. There might be other fitting responses to this situation as well, such as assuring and comforting the child, but there is no conflict between them. Reasons, by contrast, can conflict, and often do. They compete in the overall determination of how to respond.

That reasons compete is a key feature of normative reasons (Schroeder 2021).⁶ To illustrate this in the pond example, there are reasons that count in favour of rescuing the child – above all, that the child will drown if not rescued. There are also reasons, albeit far less weighty ones, that count against – such as that, by not jumping in, you'll avoid getting wet. Normative reasons are normatively relevant considerations that contribute, possibly in conflicting ways, to the overall determination of how to respond.

⁶ Schroeder (2021: 34) identifies three main features of reasons: that they are act-oriented, that they can be acted for, and that they compete. I've already pointed out that both fittingness and reasons may be act-oriented, in the sense that they may both support actions, so act-orientation is not a main difference, in my view. I'll come back to this point in section 5 below. Can only reasons can be responded to, or acted for? I don't think that's quite right, but don't have the space to discuss this. In Peter (2023) I outline a way of thinking about responding to fittingness, involving what I call moral affordances.

This difference between normative reasons and fittingness has implications for their normative reach. Because it's an overall notion, fittingness captures a particularist concern. It picks out the fitting response (or the set of fitting responses) in a given situation. Normative reasons, by contrast, are considerations that apply across a range of circumstances. Their normativity is one of general principles. For example, that this helps your friend is a normative reason that is relevant both in a standard benevolence case and in a standard corruption case. In the latter, it will hopefully be outweighed by reasons appealing to honesty and fairness. Fittingness, by contrast, because it's an overall notion, is more immediately tethered to an object or a situation. Helping your friend may be fitting in a standard benevolence case, but it's not fitting in a standard corruption case.

Given this difference between reasons and fittingness, how should we understand the relation between them? As mentioned in the introduction, my aim in this paper is to raise a problem for the claim that reasons are the mark of the normative that arises if we accept a fittingness-based account of normative reasons. That is to say, my question is, what are the consequences for the normative claim if we accept the fittingness-priority view of the structure of the normative realm? According to the fittingness-priority view, fittingness is explanatorily prior to normative reasons: normative reasons are (at least partially) grounded in fittingness, and not the other way around.⁷

⁷ As I won't have anything to say in this paper about the relation between fittingness and further normative properties or relations, I don't need to commit to a fittingness-first view of the structure of normativity. My focus is just on the relation between fittingness and reasons. Cullity (2022) uses similar terminology.

3. McHugh and Way's Reasoning-account of Reasons

Given the fittingness-priority view, how should the relationship between fittingness and reasons be understood? A number of different proposals have been put forward. Some take reasons to explain the fittingness of a response (e.g. Chappell 2012), others take it as evidence for the fittingness of that response (e.g. Thomson 2008; Sharadin 2015). A particularly promising proposal is a reasoning account of reasons that is anchored in fittingness (e.g. McHugh and Way 2016, 2022). A reasoning-account is promising because it sheds light on how reasons mesh with reasoning. A reason, it seems, is “an item in (actual or possible) reasoning” (Hieronimy 2013: 115). And normative reasons, specifically, are the considerations that should influence our reasoning when we’re reasoning well. As Mark Schroeder (2009: 26) puts it, when an agent “is reasoning well, the kinds of thing about which he should be thinking are his reasons”. This feature of normative reasons needs accounting for.⁸ McHugh and Way’s version of a fittingness-priority view offers us a good handle on this. Their reasoning account is also one of the most developed fittingness-based accounts of normative reasons, which provides a further rationale for dwelling on it here.

McHugh and Way take their reasoning account to offer a constitutive account of reasons (2022: 10ff). On this account, what it is for a consideration to be a normative reason is to be explained in terms of fittingness. Specifically, normative reasons are propositions that figure as premises in patterns of good reasoning that start from fitting attitudes. As they put it (2022: 27): “For $\langle p \rangle$ to be a reason for A is for $\langle p \rangle$ to be a premise of good reasoning from fitting responses to A.” Good reasoning, they define as a pattern of reasoning that is fittingness-preserving (2022: 42).⁹ That is to

⁸ I develop this point in Peter (2019).

⁹ A pattern of reasoning – such as modus ponens, for example – isn’t just a one-off instance of reasoning (McHugh and Way 2022: 43).

say, good reasoning is a pattern of reasoning that, if it starts from fitting premises p_1, \dots, p_n , reaches a conclusion that is fitting, too.

Consider the proposal in the context of a means-end reasoning pattern (McHugh and Weigh 2022: 41). Suppose you intend an end E that is fitting (first premise). You also fittingly (i.e. truly) believe that M is a necessary means for E. Your conclusion that you should intend M is then also fitting. Because this is good reasoning, the premises of this reasoning pattern are normative reasons. That E is a fitting end, and that M is necessary for E are both normative reasons to intend M in this case. To use one of their own illustrations of the proposal (2016: 587), that disabling the weapon of a person who is out-of-control prevents them from killing others is a reason to intend to disable the weapon, if you fittingly intend to prevent them from killing people and fittingly believe that disabling the weapon is necessary to achieve this. What makes it the case that disabling the weapon to prevent further killings is a normative reason is that it's a premise in a good pattern of reasoning that starts from a fitting response – preventing the person from killing others – that would allow you to form the intention to disable the weapon, which would be a further fitting response.

How well suited is this proposal to account for the difference between normative reasons and fittingness that I identified earlier? Their proposal is consistent with treating fittingness as an overall notion. Fittingness picks out a response that is overall merited or appropriate, not just a consideration that counts in favour of a response.¹⁰ Their proposal struggles to give a good account

¹⁰ They waver a bit on what that sort of relation fittingness is. For intentions, they sometimes identify it with permissibility (McHugh and Way 2018a) and elsewhere with choice-worthiness (McHugh and Way 2018b, 2022). My preferred interpretation is that fittingness is a kind of ought, but not the ought of obligation – I

of how reasons compete, however. A first pass argument is this. If reasons are premises in patterns of good reasoning, and if good reasoning is anchored in a fitting response, and fittingness is an overall notion, then there can't be a conflict in the starting-point of good reasoning. But that means that such patterns of good reasoning will only yield reasons that don't compete. Consider the issue in the pond example. In that context, as we saw, you have a normative reason to rescue the child and rescuing the child is also the fitting response. Their account of reasons is consistent with that. There is a good pattern of reasoning that gets you from fittingly intending to prevent the child from drowning via fittingly believing that the necessary means is to rescue the child to the fitting intention to rescue the child. However, as I highlighted earlier, you also have a competing reason, albeit a much less weighty one, not to rescue the child in this situation – that you'll avoid getting wet. That reason is also normative. It picks out a further consideration that may influence your overall assessment of how to respond. It's outweighed by the reason to rescue the child, of course. But the fact that it's outweighed by the reason to rescue the child doesn't undercut its status as a normative reason that applies to this situation. It's a competing normative reason that needs accounting for.

The problem this creates for McHugh and Way's proposal is that not wanting to get wet is not a fitting response to this situation. But if it's not a fitting response, then it can't be a starting-point for a good pattern of reasoning that would involve that you'll avoid getting wet as a premise. So their analysis can't identify that you'll avoid getting wet as a competing normative reason that applies to this situation. Something has been lost that is distinctive of the reasons relation.

expand on this in Peter (2023). This interpretation is in line with Ewing (1939), for example. Maguire (2018: 780) also describes fittingness (in relation to attitudes) as a kind of ought.

This first-pass argument might be too quick, however. McHugh and Weigh can respond to my objection by pointing out how they might accommodate conflicting reasons. Building on Way (2017), they argue that their account can accommodate outweighed reasons on the basis of the defeasibility of reasoning.¹¹ Even good patterns of reasoning can be defeated. An otherwise good pattern of reasoning may turn into a bad pattern of reasoning if further premises are added. For example, intending to meet your friend for dinner given that you promised to meet them may be the result of a good pattern of reasoning. Everything else equal, that you promised to meet them is thus a reason to intend to meet them. But suppose you've just been made aware that a close relative is in hospital and needs your support. If you were still to conclude that you should meet your friend, even after adding this premise, then this would not be supported by good reasoning. Good reasoning now yields a reason to break your promise. Given the additional premise, the reason to keep your promise has been defeated. But that it was originally part of a good pattern of reasoning explains why it's a normative reason. And the more comprehensive reasoning that involves the additional premise explains why the reason to keep your promise is outweighed by the reason to support your relative.

This solution goes some way to alleviate the worry. I think they're right about the defeasibility of reasoning and about its bearing on the weight of reasons. The defeasibility of even good patterns of reasoning implies that reasons, analysed as premises in such reasoning, may be outweighed. I think this is a persuasive account of the weight of reasons. The solution doesn't fully address the problem I've highlighted, however. And that's because their solution only works if we're presupposing that the reasons that are involved in a defeated pattern of reasoning are normative. But the normativity of potentially conflicting reasons needs accounting for and that creates the problem.

¹¹ They provide this account in chapter 5 of McHugh and Way (2022).

Recall that their claim is that reasons are normative in virtue of being premises in patterns of reasoning that are fittingness-preserving. To secure the normativity of reasons, they argue, it's not enough to appeal to well-established patterns of reasoning such as means-end reasoning. Not any hypothetical reasoning will do. Patterns of reasoning only yield normative reasons, on their account, if the reasoning starts from fitting attitudes (McHugh and Way 2022: 42ff).

If we focus on this normative claim, we can see that pointing to the defeasibility of reasoning is inadequate to account for competing normative reasons. Consider the issue in the pond example again. In the pond case, not wanting to get wet is not a fitting starting-point of reasoning. It's not just that the reason to stay dry is defeated by a better pattern of reasoning. There is, in this situation, no good pattern of reasoning that starts from the intention of staying dry because intending to stay dry is not a fitting response to this situation. So their proposal can't explain why that you'll avoid getting wet is a competing normative reason in this situation. It's not a normative reason at all, on their analysis. But that can't be right.¹²

To put the point more generally, if normative reasons inherit their normativity from reasoning that takes as its starting-point a fitting response to a given situation, and if fitting responses to a given situation can't conflict, because fittingness is an overall notion, then the normative reasons that apply in a given situation can't conflict. But that implies that McHugh and Way's analysis can't account for competing normative reasons.

¹² McHugh and Way have a discussion of what they call "the problem of bad starting-points" (2022: 33). They allow that there is good reasoning from bad-starting points, and they allow that it can generate subjective reasons. My point is that it can't generate objective normative reasons.

4. A Linking Principle

To account for the tendency of normative reasons to compete, we need an account of normative reasons that is anchored in fittingness but that can accommodate possible conflicts among those reasons. While I've argued that McHugh and Way's reasoning account of normative reasons fails in this regard, I think that a proposal in the vicinity of theirs can succeed. To account for the tendency of normative reasons to compete, we must rethink the set of admissible starting-points in good reasoning. In this section, I put forward a principle that sheds light on the link between fitting responses and normative reasons.

Before I proceed to spell out this linking principle, let me add some clarifications. First, like McHugh and Way (2022), I take normative reasons to be propositions (Peter 2019). The motivation for this way of understanding reasons is that an account of reasons must explain how reasons can figure in deliberation. Reasons understood as propositions meet the deliberative constraint because propositions are the shareable content of considerations. Relevant propositions are considerations that favour some response.

Second, what exactly is the link between fitting starting-points of reasoning and normative reasons? Put in the context of McHugh and Way's account, what are admissible premises in the kind of reasoning that is constitutive of normative reasons, and what makes those premises admissible? As we saw, not all premises in a good pattern of reasoning are normative reasons. Only premises in a good pattern of reasoning that starts from fitting responses are normative reasons. We thus need to clarify the link between fitting responses to a situation and normative reasons.

My proposal is that admissible premises are, or are derivative of, true propositional representations of the facts that make a response fitting. For example, that the child is drowning is a true representation of the fact that makes rescuing the child fitting. Similarly, that it prevents further killings is a true representation of the fact that makes disabling the weapon is fitting.

To explain this point further, in an earlier paper, I argued that normative reasons are true propositions that represent a certain kind of favouring in thought (Peter 2019). I developed the idea drawing on Christine Korsgaard's distinction between acting in response to a fact and acting in response to the description of this fact as a reason for action (Korsgaard 2008). Consider her example of the lioness: while a lioness may respond to a fact (for example, by bringing her cubs to safety), she's probably not responding to the representation of that fact – that there is a predator nearby, say – as a reason for action. The earlier paper left it open what the normative facts are that are represented by normative reasons. I can now make my earlier proposal more specific by explaining (at least some) normative reasons as propositions that represent fit-making facts in thought.

A third clarification concerns the scope of the linking principle that I'm about to introduce. I understand the linking principle as making a partial grounding claim: the representation of fit-making facts plays a role in the explanation of what it is to be a normative reason. The linking principle only covers normative reasons that are true propositional representations of fit-making facts. Unlike McHugh and Way (2022), I'm not attempting to give a complete account of what constitutes normative reasons, and I accept that something like their reasoning account is necessary to complete the picture. I accept, for example, that a reasoning account is necessary to shed light on relations among normative reasons. In particular, I find their account of the weight of reasons quite persuasive.

To spell out the linking principle, it helps to make use of my earlier characterisations of the fittingness and the reasons relations. The latter, following Scanlon (2014: 31), is a relation $R(p, a, x, c)$ that picks out a reason p for agent x in some circumstances c to do a . The former is a relation $F(a, x, c)$ that picks out the fitting response a (an action or attitude) for x in some circumstances c . In a first approximation, my proposal is that we understand the link between the two as follows:

Linking principle (first approximation): for some normative reasons, for p to be a normative reason for x to do a in circumstances c is for p to truly represent in thought a fact that makes a the fitting response for x in c .

To explain, a fitting response a is a doing or a feeling, which stands in an appropriate relation to the circumstances c . That relation can be represented in thought in a proposition that captures the way in which relevant aspects of the circumstances – the fit-making facts – support the fitting response. If it's a true representation of a fit-making fact, then the proposition p is a normative reason – it's a consideration that favours that particular response.¹³ In the pond case, assuming that rescuing the child is the fitting response, and that the proposition that the child is drowning is a true representation of the fact that makes rescuing the child fitting, then that the child is drowning is a normative reason to rescue the child. In general, this first version of the principle explains the link between the fittingness relation $F(a, x, c)$ and the reasons relation $R(p, a, x, c)$ as follows: while the – normatively more fundamental – fittingness relation identifies the fitting response a for an agent x in a given situation c , the reasons relation then specifies a consideration p that is a true

¹³ If the representation is false, then proposition p is merely an apparent reason (Parfit 2011).

representation of a fact – a feature or aspect of c – that makes a the fitting response for x in this situation.

The linking principle focuses on normative reasons that directly represent fit-making facts in thought. It doesn't say anything about normative reasons that are only indirectly related to the representation of fit-making facts, say because they are derivative of normative reasons that directly represent fit-making facts. That's why it only holds for some normative reasons.¹⁴

The linking principle in this first approximation sheds some light on the question of what is an admissible starting-point for good reasoning. But it's too narrow to account for competing reasons. As we saw, in the pond example, there isn't just a normative reason to rescue the child. There is also a – much less weighty – normative reason that favours not to do so that needs accounting for. To accommodate competing reasons, the link between fitting responses and normative reasons can't therefore be too tight – contrary to what we get from McHugh and Way's proposal (2016, 2022), as discussed above.

To allow for competing reasons, we must reconcile the particularity of the fittingness relation with the tendency of normative reasons to apply across situations. As mentioned in section 2, a normative reason is a consideration that isn't just specific to a particular situation. It applies across a number of situations and then competes with other reasons that apply in a given situation. This creates a tension between fittingness, which is an overall notion, and normative reasons, which isn't sufficiently acknowledged in McHugh and Way's analysis, and which presents a key challenge for a fittingness-based account of reasons.

¹⁴ I'll elaborate on this point in section 5, where I offer the final statement of the linking principle.

We can accommodate these concerns by keeping the fittingness relation tethered to some particular circumstances c while linking normative reasons to fitting responses in situations s that overlap with c . The linking principle in its second approximation captures this idea:

Linking principle (second approximation): for some normative reasons, for p to be a normative reason for x to do a in circumstances c is for p to truly represent in thought a fit-making fact in c or in some situation s that overlaps with c .¹⁵

The idea here is this. As per the first statement of the linking principle, suppose p_c is the consideration that truly represents in thought a fact that makes a the fitting response in circumstances c (e.g. that the child is drowning). p_c is then a normative reason for x to do a (jump in to rescue the child). But p_c doesn't exhaustively describe the circumstances c . Other true thoughts about c are possible. And the contents of some of those thoughts will be propositions that truly represent fit-making facts for some response b in different situations s . As per the original statement of the linking principle, these propositions p_s are normative reasons that apply in situations s . However, because these propositions are also true in c , those situations s overlap with c . The linking principle says that relevant considerations p_s are normative reasons that apply in c , too. They represent favouring considerations that bear not just on how to respond in s , but also on how to respond in c . So a normative reason that applies to a given situation c is a proposition that favours some response a by x in c and that is a true representation of a fit-making fact in either c or in a situation s that overlaps with c .

¹⁵ While this is a near final statement, the final statement of the linking principle will follow in section 5. I'm grateful to Hrafn Asgeirsson for pointing out a problem with an earlier version of this statement.

To illustrate, the pond case overlaps with a situation where you're also at risk of getting wet.

Suppose you're considering leaving the umbrella at home, but taking the umbrella would be fitting.

That you'll avoid getting wet (p_s) on your way to the office is a normative reason to take the umbrella and that's because it truly represents the fit-making fact of taking the umbrella in this situation. Because it's a true representation of the fit-making fact explains why that you'll avoid getting wet is a normative reason for doing the thing that keeps you dry in this situation.

According to the linking principle, because that you'll avoid getting wet is a normative reason in the umbrella case, it's also a normative reason in any situation where the risk of getting wet bears on your response, such as the pond case. That you'll avoid getting wet (p_s) is a true proposition in the pond case that supports not rescuing the child, and this consideration inherits its normativity from situations in which not getting wet is fitting. The linking principle thus explains why you not only have a reason to rescue the child in the pond case, which is also the fitting response in this case, you also have a competing, though much weaker, reason not to jump into the pond.¹⁶

By contrast, that the liquid in the pond includes H₂O, say, is probably not the sort of proposition that constitutes a normative reason that applies in the pond case. It may be true, and it may be a fit-making fact for some response b in some situations s , but unless it favours some response a in this case, it's not a normative reason that applies. I, for one, can't think of a response that would be supported by this consideration. Conversely, that you'll save the world from evil by not rescuing the

¹⁶ As mentioned earlier, the reasoning account is helpful to account for the fact that this reason is outweighed by the reason you have to rescue the child – that the child is drowning, or that the child needs rescuing.

child is not a reason to let the child drown because, let's assume, it's not a true proposition about c (and probably not a true representation of a fit-making fact in any situation s).

Note that the linking principle doesn't just apply to reasons for action; it also applies to normative support for attitudes. Consider a discrimination case, in which being angry is the fitting response – say someone (you or someone else) has been unjustly ignored. If that someone has been unjustly ignored truly represents the fittingness of an angry response, then that someone has been unjustly ignored is a normative reason for getting angry in this case (p_c). Other reasons might apply in that case, too. For example, suppose that in some cases that overlap with the discrimination case, forgiving is a fitting response. And if that there has been an unintentional transgression is the relevant fit-making fact, the representation of that fact (p_s) is then a reason for forgiving in the discrimination case as well.

Some of those considerations p_s might favour the same response as p_c , although on different grounds and perhaps only in part. In the pond example, that you enjoy swimming might be a supporting reason for rescuing the child. That reason favours jumping in, even if it's not out of care for the child. It draws its normativity from situations in which seeking pleasure from swimming is fitting. But the normative reasons p_s that apply in a given situation (e.g. that you'll avoid getting wet, that there has been an unintentional transgression, etc.) will often support responses other than the fitting response in c . p_c and p_s are then competing reasons that apply in c . The linking principle thus explains how a fittingness-based account of normative reasons can accommodate the tendency of reasons to compete.

5. The Normativity of Fittingness

Later in this paper, I will use the linking principle to show that there are two kinds of normative reasons – RKR (right kind of reasons) and runaway reasons. Before continuing with my main argument, however, I should pause and consider a possible objection to the linking principle (and any fittingness-based account of normative reasons). The objection is that because fittingness is only thinly normative, not robustly normative, anchoring an account of reasons in fittingness can't secure their normativity.

As I understand the distinction, robust normativity involves normative facts that exert some authority or prescriptivity while thin normativity does not.¹⁷ Why might fittingness not be robustly normative? The answer depends on the internal standard view of fittingness that is currently dominant. On this view, the fittingness of a response is determined by an internal standard for this response – typically a constitutive standard or some other standard of correctness. For example, on the internal standard view, the fittingness of a response such as admiration depends on the constitutive standard for admiring, or on some other standard of correctness specific to admiration. The problem that this view creates is that internal standards need not be robustly normative. For example, there are etiquette-type standards for how to hold a fork at a fancy dinner party, and there are rules for how to correctly move chess pieces in a game of chess. If such standards are normative at all, they are only thinly, or formally, so. Internal standards come cheap (Lord 2023: 262) and their normativity is not robust enough to anchor normative reasons.¹⁸

¹⁷ For discussions of the distinction between thin or formal normativity and robust or authoritative normativity, see McPherson (2018) and Côté-Bouchard and Littlejohn (2018).

¹⁸ Lord and Sylvan (2019) and Rowland (2022) discuss worries about the normativity of standards of fittingness.

My reply to this objection is that I accept it given an internal standard view of fittingness. But I reject the internal standard view in favour of a relational view of fittingness. On the relational view, the fittingness of a response depends on how the response relates to the situation overall, and not (just) on the response-type that it is. The fitting response is the one that is called for from a given agent by a given situation.

The internal standard view of fittingness, while currently dominant, appears to be a relatively recent development.¹⁹ The relational view used to be more common. We find this view of fittingness articulated in the passage from Broad that I quoted above. To repeat, “[fi]ttingness or unfittingness is a direct ethical relation between an action or emotion and the total course of events in which it takes place” (Broad 1930 / 2014: 219). We also find it in A. C. Ewing’s work, who distinguishes between two types of ought – the ought of fittingness and the ought of moral obligations. With regard to the first type of ought, he writes, “‘the act I ought to perform’ stands for the act which is most fitting or most desirable in view of the situation” (1939: 3). So, again, the idea is that the fittingness of a response depends on the overall situation, and not (just) on the response-type. More recently, Howard and Leary (2022) also reject the internal standard view of fittingness in favour of what looks to me like a relational view.

Historical and contemporary precedents apart, why adopt a relational view of fittingness? While there is lots to say in response to this question, I lack the space to fully answer it here. But the short version of the answer is that the relational view is independently attractive, and it secures the robust normativity of the fittingness relation. One reason it’s independently attractive is that it can

¹⁹ On my understanding, Schroeder (2010) was quite influential in this regard. See also Howard and Leary (2022) for a critical discussion of this development; my argument here parallels theirs in many respects.

account for the fittingness of both actions and attitudes, as we saw in the historical examples. How does the relational view secure robust normativity? The problem for the internal standard view of fittingness arises because response-types don't carry any authoritative or prescriptive weight. Norms for how to play chess correctly, norms of dinner party etiquette, or norms that distinguish an envious response from an angry response might thus better be called norms of individuation – suitable for differentiating between different types of actions or attitudes.

Such standards of individuation say nothing about what the right response overall is to a given situation. As we might also put it, they are non-strict (Maguire 2018). Should you be playing chess or should you be taking care of your child instead? Is social conformity called for? Should you be feeling envious, or would it be more appropriate to be happy for a friend who won some award? The relational view of fittingness aims to give an answer to these questions (although the details have to be worked out elsewhere). As such, it's relevantly authoritative or prescriptive. Because the internal standard view does not, it would be a mistake to identify the standard of fittingness with norms of individuation if fittingness is to play a fundamental role in a theory of normativity. A fittingness-priority view of the relation between fittingness and normative reasons – let alone a fittingness-first view of normativity – is promising only if fittingness is robustly normative, and only the relational view can secure this.

If we plug the relational view of fittingness into the linking principle, this sheds further light on what normative reasons represent in thought. It's not internal standards of fittingness, at least not if the normative reason in question is robustly normative. A robustly normative reason that applies in c is a true proposition that favours some response a and that is a true representation of a fit-making fact in either c or in a situation s that overlaps with c . My final statement of the linking principle is thus:

Linking principle (final statement): for p to be a robustly normative reason for x to do a in circumstances c is for p to truly represent in thought a fit-making fact in c or in some situation s that overlaps with c .

6. The Possibility of Runaway Reasons

In this section and the next, I focus on an implication of my account of robustly normative reasons. The implication is that there are robustly normative considerations that are grounded in facts about fittingness, but that favour responses other than the response that is fitting in a given situation. I call such reasons runaway reasons. In this section I explain what runaway reasons are. In the next section, I explain the problem that runaway reasons give rise to.

To recap, I argued that a fittingness-based account of reasons can't rest on too tight a link between fitting responses and reasons, or it loses the capacity to account for competing reasons. I took it as a given that normative reasons can compete, and that a fittingness-based account of reasons must accommodate this feature. The linking principle that I propose accommodates this feature. It does so by allowing that true representations of fit-making facts in situations that overlap with the circumstances under consideration are normative reasons that also apply in those circumstances. All those normative reasons are robustly normative.

All normative reasons contribute to determining what you have most reason to do in a given situation. But they do not all favour the fitting response, and some only do so partially. If we allow that the set of fit-making facts that anchor normative reasons isn't limited to the fit-making facts in the circumstances under consideration, as the linking principle does, then not all normative reasons that apply in a given situation, even if they are robustly normative, will track the fitting response. This shows that there is a normatively significant distinction among the normative reasons that

apply in a given situation. Only some pick out the fitting response to a given situation – that’s consideration p_c . Others do not – those are the competing considerations, and the considerations that only partially support the fitting response, p_s . I call the reasons p_s that do not pick out the fitting response runaway reasons. A runaway reason is a robustly normative reason that applies in a given situation, but that doesn’t (fully) support the fitting response in this situation.

Runaway reasons are a result of the differences in the normativity of fittingness and reasons.

Fittingness, as explained earlier, is an overall notion that picks out a fitting response to a given situation. The relational view of fittingness that I introduced in the last section sheds further light on this. Normative reasons, by contrast, are contributory and they are not normatively confined to a given situation. They pick out considerations for or against some response, which may or may not be the fitting response in this situation. On a fittingness-based account of reasons, the normativity of reasons is anchored in fittingness. But to accommodate the tendency of reasons to compete, the domain of normative reasons that apply in a given situation can’t just be limited to the fitting response to a given situation. And that creates a tension between the normativity of fittingness and the normativity of reasons.

Deliberation can help keeping runaway reasons in check. In reasoning about how to respond based on those considerations, if we reason well, we aim at fitting responses. We weigh-up the normative reasons that apply to come to a conclusion about how to respond. Defeasible reasoning, in particular, as described by McHugh and Way (2022), can help identify outweighed reasons. If all goes well, the deliberation will pick out the fitting response – for reason p_c and perhaps additional supporting reasons p_s . In the pond case, if the consideration that the child is drowning is taken into account alongside the consideration that, by not jumping, you’ll avoid getting wet, it becomes apparent that the reason you have to rescue the child outweighs the reason not to.

But all may not go well. In some cases, runaway reasons can support responses other than the fitting response. What factors could cause such an outcome? Assuming that deliberation otherwise functions well – that it's well-intended, careful, etc. – a first factor is whether p_c is available as a normative reason at all. In circumstances in which the fitting response to a given situation and the fit-making facts remain uncharted territory, p_c can't be a normative reason that applies to this situation. For example, imagine a case of sexist discrimination before sexism has been represented as a cause of unfitting responses.²⁰ Non-discrimination would be fitting in that case, but that the attitude or action is sexist wouldn't be available as a reason to support non-discrimination. If p_c isn't available as a normative reason, it can't keep runaway reasons – normative reasons that support responses other than the fitting response – in check.

A second factor is, how weighty are the normative reasons that have their origin in situations that overlap with c ? Consider a Kamm-type thought experiment, which focuses on a situation s in which, instead of saving the drowning child, a large number of headaches could be cured instead. While the fitting response in the pond example, let's assume, is to rescue the child, deliberation might lead us to conclude that the normative reason to rescue the child is outweighed by the reason to cure the headaches. That you can cure a very large number of headaches is a runaway reason in this case, but one that may be weightier than the reason that favours the fitting response.

To illustrate the idea of runaway reasons further, take a hiring case. A university department has been reviewing applications and interviewing candidates. There is now deliberation on whom to offer the job to. A broad range of reasons in favour of different candidates is considered in

²⁰ This example is loosely based on Harman (2011).

deliberation, let's assume all with the aim to identify the best candidate (the fitting response). Reasons considered in support of different candidates include that A has the strongest research record, that B is a great teacher, that C would thrive most in this department, that D has a great grasp of the history of philosophy, etc. Suppose that appointing candidate C is the fitting response in this case, and that they would thrive most is the normative reason that truly represents the fittingness relation that obtains (p_c). But the department can't just read p_c off the situation. To identify p_c , if all goes well, they must consider all the normative reasons that apply (which include p_c and a range of potentially competing reasons p_s) and weigh them. In doing so, there is always the risk that even well-intentioned deliberation gets hijacked by runaway reasons. The risk is of course even bigger if the deliberation is not well-intentioned, or if it's subject to biases that eclipse the fitting response and the normative reason(s) that favour it.

In sum, in this section I've argued that one upshot of the linking principle is that there are robustly normative reasons that support responses other than the fitting response. This is not a bug of the linking principle, I contend. Instead, it's an implication of a fittingness-based account of competing normative reasons, and an underappreciated feature of our normative lives. In the next section, I explain how the possibility of runaway reasons creates a problem for the claim that normative reasons are the mark of the normative.

7. The Runaway Reasons Problem

Recall that I characterised reasons-first views as committed to two independent claims, an explanatory claim and a normative claim. The WKR problem highlights that a view that takes normative reasons to be explanatorily first can fail to appropriately track other normative properties or relations. A fittingness-based solution to the WKR problem abandons the claim that reasons are explanatorily first. Can this solution support a revised version of the normative claim, namely that

normative reasons, as long as they are grounded in fittingness, are the mark of the normative? The answer to this question is no, and that's because of runaway reasons.

Runaway reasons, as we saw, are robustly normative. Their normativity derives from the normativity of fittingness. Yet they might not track the fitting response (and perhaps other normative properties and relations) in a given situation. Not all robustly normative reasons are RKR. Runaway reasons, although robustly normative because anchored in fittingness, are not RKR.

The possibility of runaway reasons confronts us with the challenge of distinguishing between normative reasons that are RKR – specifically, those that track the fitting response in a given situation – and normative reasons that are runaway reasons. This runaway reasons problem compounds the WKR problem by highlighting that not all fittingness-based reasons are RKR. Some normative reasons, even though grounded in fittingness, are runaway reasons and may lead us normatively astray. This puts pressure on the normative claim, the claim that normative reasons, at least as long as they are fittingness-based, are the mark of the normative.

As we might also describe it, the runaway reasons problem implies that not all normative concerns can be dissolved in the medium of normative reasons without residue. But the runaway reasons problem is not the only possible source of normative residues, and we need to distinguish the runaway reasons problem from these other cases. In those other cases, normative residues arise from a tension among irreconcilable substantive concerns, not from a tension between fittingness and reasons, unlike with runaway reasons.²¹ What Bernard Williams (1981) calls agent-regret is a prime example. Agent-regret is a feeling of distress in the aftermath of an event that was caused by

²¹ I'm grateful to Ezequiel Monti, Debbie Roberts, and Michael Smith for drawing my attention to this issue.

the agent, but for which the agent is not morally responsible. To illustrate, consider a mother who is working to support her family but who feels sad about leaving her child for the day.

First, note that such cases might have their origin in what is fitting in a given situation, and not involve normative reasons at all. To see this, recall that I've allowed that more than one response might be fitting in a given situation, compatible with fittingness being an overall notion. In the case of the mother, less assume that both her choice of action – to go to work to support her family – and her feeling of sadness are overall fitting responses to her situation. Her action of going to work and leaving her child for the day will leave a normative residue in the form of her sadness. But no runaway reasons are involved here.

Second, the tension might reproduce at the level of normative reasons, but without involving runaway reasons. To illustrate the point again in the case of the mother, given that both her going to work and her experience of sadness are fitting, there are normative reasons that arise from the representation of the fittingness of those responses. That it supports her family is a normative reason for her to go to work. That she has to leave her child for the day is a normative reason for her to feel sad. Again, there are no runaway reasons involved at this point.

This said, runaway reasons are possible in such scenarios, too. Suppose that, reflecting on her situation, the mother considers her reasons to go to work (e.g. that it supports her family) and reasons not to go to work (e.g. that she'll avoid feeling sad, or that she'll avoid leaving her child for the day). In the scenario as I've described it earlier, the reasons she has not to go to work are runaway reasons. They are normative reasons that applies in this scenario – because there are situations that overlap with hers in which avoiding feeling sad, or avoiding leaving her child for the

day, would be fit-making facts, and they support not going to work. But they are not normative reasons that pick out a fitting response in this scenario, and so they are runaway reasons.

As we might also put it, the runaway reasons problem that I have identified arises because the set of robustly normative, fittingness-based reasons is larger than the set of RKR. In this sense, the problem is the inverse of the RKR problem that Erol Lord and Kurt Sylvan identify (2019). The RKR problem is that not all (fittingness-based) RKR reasons are robustly normative – some are only thinly normative. The RKR problem trades on the claim, which I've rejected, that fittingness isn't robustly normative. The runaway reasons problem that I'm describing is that while all fittingness-based reasons are robustly normative, not all of them are RKR – some are runaway reasons. The runaway reasons problem deepens the WKR problem. It shows that a fittingness-based account of normative reasons doesn't guarantee RKR. Even fittingness-based reasons may be runaway reasons.

The runaway reasons problem implies that the normativity of reasons and the normativity of fittingness can be in tension. What we have most reason to do may not be the fitting response. Runaway reasons thus point to a possible bifurcation in our normative concerns – between responding fittingly and responding in a way that can be supported by normative reasons. And the consequence of this bifurcation is that normative reasons are not the sole mark of the normative.

8. Concluding Remarks

On fittingness-priority views, the normativity of reasons is anchored in the normativity of fittingness. Many think that such a fittingness-priority view is compatible with taking normative reasons, as long as they are anchored in fittingness, as the mark of the normative. In this paper, I revealed a problem for this view. The problem is that some fittingness-based reasons are runaway reasons. Runaway

reasons are robustly normative, but they support a response that is not the fitting response in given circumstances. Runway reasons thus have the potential to lead us normatively astray.

Acknowledgements

For helpful feedback and suggestions on earlier versions of this paper I'm most grateful to Sara Gorea, Ulrike Heuer, Ezequiel Monti, David O'Brien, Debbie Roberts, Mark Schroeder, Michael Smith, and audiences at The Madison Metaethics Workshop, the Oxford Jurisprudence Discussion Group, and the Trident Seminar series in legal philosophy.

References

- Broad, C.D. 1930 / 2014. *Five Types of Ethical Theory*. London: Routledge.
- Chappell, Richard (2012) 'Fittingness: The Sole Normative Primitive', *The Philosophical Quarterly*, 62(249), pp. 684–704.
- Côté-Bouchard, Charles and Littlejohn, Clayton (2018) 'Knowledge, Reasons, and Errors About Error Theory', in C. Kyriacou and R. McKenna (eds) *Metaepistemology: Realism and Anti-Realism*. Cham: Springer International Publishing (Palgrave Innovations in Philosophy), pp. 147–171.
- Cullity, Garrett (2022) 'Reasons and Fit', in C. Howard and R.A. Rowland (eds) *Fittingness: Essays in the Philosophy of Normativity*. Oxford University Press, p. 151 – 175.
- Dancy, Jonathan (2006) 'What Do Reasons Do?', in *Metaethics after Moore*. Oxford: Oxford University Press.
- Ewing, A.C. (1939) 'A Suggested Non-Naturalistic Analysis of Good', *Mind*, 48(189), pp. 1–22.
- Harman, Elizabeth (2011) 'Does Moral Ignorance Exculpate?', *Ratio*, 24(4), pp. 443–468.
- Hieronymi, Pamela (2013) 'The Use of Reasons in Thought (and the Use of Earmarks in Arguments)', *Ethics*, 124(1), pp. 114–127.

- Howard, Christopher (2016) 'In Defense of the Wrong Kind of Reason', *Thought: A Journal of Philosophy*, 5(1), pp. 53–62.
- Howard, Christopher (2018) 'Fittingness', *Philosophy Compass*, 13(11), p. e12542. Available at: <https://doi.org/10.1111/phc3.12542>.
- Howard, Christopher (2019) 'The Fundamentality of Fit', *Oxford Studies in Metaethics*, 14. **
- Howard, Christopher and Leary, Stephanie (2022) 'In Defence of the Right Kind of Reason', in Chris Howard and R.A. Rowland (eds) *Fittingness: Essays in the Philosophy of Normativity*. Oxford University Press, pp. 221 – 241.
- Korsgaard, Christine M. (2008) 'Acting for a Reason', in C.M. Korsgaard (ed.) *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford University Press, p. 207 - 229.
- Lord, Erol (2023) 'Everything First', *Aristotelian Society Supplementary Volume*, 97(1), pp. 248–272.
- Lord, Erol and Sylvan, Kurt (2019) 'Reasons: Wrong, Right, Normative, Fundamental', *Journal of Ethics and Social Philosophy*, 15(1).
- Maguire, Barry (2018) 'There Are No Reasons for Affective Attitudes', *Mind*, 127(507), pp. 779–805.
- McHugh, Conor and Way, Jonathan (2016) 'Fittingness First', *Ethics*, 126(3), pp. 575–606.
- McHugh, Conor and Way, Jonathan (2018a) 'What is Good Reasoning?', *Philosophy and Phenomenological Research*, 96(1), pp. 153–174.
- McHugh, Conor and Way, Jonathan (2018b) 'What is Reasoning?', *Mind*, 127(505), pp. 167–196.
- McHugh, Conor and Way, Jonathan (2022) *Getting Things Right: Fittingness, Reasons, and Value*. Oxford University Press.
- McPherson, Tristram (2018) 'Authoritatively Normative Concepts', in R. Shafer-Landau (ed.) *Oxford Studies in Metaethics 13*. Oxford University Press.
- Parfit, Derek (2011) *On What Matters: Volume One*. Oxford, New York: Oxford University Press (The Berkeley Tanner Lectures).

- Peter, Fabienne (2019) 'Normative Facts and Reasons', *Proceedings of the Aristotelian Society*, 119(1), pp. 53–75.
- Peter, Fabienne (2023) 'Moral affordances and the demands of fittingness', *Philosophical Psychology*, online first. Available at: <https://doi.org/10.1080/09515089.2023.2236120>.
- Rabinowicz, Wlodek and Rønnow-Rasmussen, Toni (2004) 'The Strike of the Demon: On Fitting Pro-attitudes and Value', *Ethics*, 114(3), pp. 391–423.
- Raz, Joseph (1999) *Engaging Reason*. Oxford: Oxford University Press.
- Rowland, Richard A. (2022) 'The Authoritative Normativity of Fitting Attitudes', in R. Shafer-Landau (ed.) *Oxford Studies in Metaethics, Volume 17*. Oxford University Press, pp. 108 - 137.
- Scanlon, Thomas M. (1998) *What We Owe to Each Other*. Cambridge: Harvard University Press.
- Scanlon, Thomas M. (2014) *Being Realistic About Reasons*. Oxford: Oxford University Press.
- Schroeder, Mark (2010) 'Value and the Right Kind of Reason', *Oxford Studies in Metaethics*, 5, pp. 25–55.
- Schroeder, Mark (2009) *Slaves of the Passions*. Oxford, New York: Oxford University Press.
- Schroeder, Mark (2021) *Reasons First*. New York: Oxford University Press.
- Sharadin, Nathaniel (2015) 'Reasons Wrong and Right', *Pacific Philosophical Quarterly*, 97(3), pp. 371–399.
- Thomson, Judith J. (2008) *Normativity*. Chicago: Open Court.
- Way, Jonathan (2017) 'Reasons as Premises of Good Reasoning', *Pacific Philosophical Quarterly*, 98(2), pp. 251–270.
- Williams, Bernard (1981) *Moral Luck*. Cambridge: Cambridge University Press.