

## Favoring

Antti Kauppinen ([a.kauppinen@gmail.com](mailto:a.kauppinen@gmail.com))

Trinity College Dublin

Forthcoming in *Philosophical Studies*.

Is it possible to give an informative and illuminating non-circular characterization of what it is for something to be a reason? Thomas Scanlon, Derek Parfit, and many others deny it. Scanlon says that “[a]ny attempt to explain what it is to be a reason for something seems to me to lead back to the same idea: a consideration that counts in favor of it.” (1998, 17) Since favoring, in turn, is a matter of providing a reason for something, any attempt at analysis leads to circularity. Parfit holds a very similar view, and says that “the concept of a reason is indefinable in the sense that it cannot be helpfully explained merely by using words” (2011, 31). These are claims about the *concept* of a reason. Call the higher-order fact that fact  $r$  favors  $\phi$ -ing a *reasons fact*. The metaphysical counterpart of the conceptual claim is then that reasons facts cannot be reduced to other normative or evaluative facts. We can express both together as follows:

Reason Primitivism: The concept of a reason is indefinable, and reasons facts form a basic category of normative facts.

Some reason primitivists hold that reasons facts are altogether *sui generis*, while others regard them as reducible to non-normative facts. For my purposes, this distinction is of little importance, as my concern is with relations within the normative realm. It matters, however, that many (but not all) reason primitivists hold that other normative and evaluative concepts or facts *are* analyzable in terms of reasons. They thus endorse the following thesis:

Reason Analyzability: All other normative and evaluative concepts or facts are analyzable in terms of reasons or reducible to reasons facts.

Together, these theses support

Reasons First: Because all other normative or evaluative concepts or facts are analyzable in terms of reasons or reason facts, but not *vice versa*, “what it is to be normative is to be analyzed in terms of reasons” (Schroeder 2007a, 81).

Joseph Raz, who may not at the end of the day endorse Reasons First, nevertheless offers a very explicit formulation of the view:

The normativity of all that is normative consists in the way it is, or provides, or is otherwise related to reasons. ... So ultimately the explanation of normativity is the explanation of what it is to be a reason, and of related puzzles about reasons. (Raz 1999, 67)

I believe that Reason Primitivism and consequently Reasons First are false. To support this contention, I will present a novel analysis of reasons in terms of *oughts* governing rational agents and thinkers, or the desirability of conducting one’s thinking in a certain way, given awareness of non-normative facts. Roughly, for *r* to be a *normative* reason for A to  $\phi$  is for it to be the case that A *ought* to conduct her  $\phi$ -relevant thinking in a more or less  $\phi$ -friendly manner, given her awareness of *r*. In mechanistic terms, this is to say that the psychological mechanisms responsible for A’s potentially  $\phi$ -ing ought to be causally influenced in the direction of  $\phi$ -ing by her awareness of *r*. For *r* to be an *evaluative* reason for A to  $\phi$  is for it to be the case that it is *desirable* for A to conduct her  $\phi$ -relevant thinking in a more or less  $\phi$ -friendly manner, given her awareness of *r*.<sup>1</sup> I claim that these two theses give us an informative and extensionally adequate analysis of the two key favoring relations between facts and psychological states, and that the resulting account has certain advantages over

---

<sup>1</sup> It has become standard to express the basic contrast in the theory of reasons in terms normative reasons, on the one hand, and motivating reasons, on the other. Both what I call normative and evaluative reasons in this paper fall on the ‘normative’ side of this standard distinction.

<sup>2</sup> In their 2013 paper, Kearns and Star offer responses to related criticisms by Broome and Brunero. I lack the

competing reductionist views. I will also address a number of objections and challenges. Many of those challenges are based on what I will call the positive reason existential fallacy: the tendency to think that any good-making feature of action or its outcome provides a reason for the action. But when we focus on what makes reasons normative or evaluative, we have principled grounds for rejecting this assumption. This points to a deeper rationale for identifying facts about favoring with facts about how we ought to or how it would be desirable for us to think, given the consideration that is a reason. Roughly, doing so reveals the sense in which reasons are normative or evaluative by linking the favoring relation between a fact and an action or belief explicitly with fittingness of attitudes towards the subject on account of how she responds to the fact.

## **1. The Rivals**

Although Reason Primitivism has become a popular view, it has not gone without challenge. In this section, I will briefly discuss two rival accounts of normative reasons in terms of oughts. John Broome (2004, 2005) takes issue with Reason Analyzability, and in doing so also challenges Reason Primitivism. Broome analyses pro tanto reasons in terms of playing a pro- or con-role in a ‘weighing explanation’ of what one ought to do. A weighing explanation is an analogue of the mechanical explanation of why a pair of scales tips in one direction rather than the other: the objects on one side have a total weight higher than those on the other (Broome 2004, 36). Broome acknowledges that the analogy is imperfect, in part due to the fact that the ‘weight’ of reasons can vary depending on what other reasons are present. But he believes that sometimes we can explain why someone ought to  $\phi$  by aggregating and comparing facts pulling in opposite directions. Pro tanto reasons are just

such facts – facts that could explain why someone ought to do something while being at least potentially outweighed by other facts. As his final definition goes,

a pro tanto reason for you to  $\phi$  is a fact that plays the for- $\phi$  role in a weighing explanation of why you ought to  $\phi$ , or in a weighing explanation of why you ought not to  $\phi$ , or in a weighing explanation of why it is not the case that you ought to  $\phi$  and not the case that you ought not to  $\phi$ . (Broome 2004, 39)

It may indeed be true that reasons play such a role in a weighing explanation. But as Stephen Kearns and Daniel Star (2008) point out, it is unclear whether we can make sense of the notion of a weighing explanation, not to mention a for- or against-role, *without already assuming* understanding of what it is to count in favor of something – that is, the very notion of a pro tanto reason. Maybe the fact that an apple is tasty plays a for-buying role in a weighing explanation of why I ought to buy one. But how could I understand this if I didn't already understand what it is for a fact to favor my doing something? We don't make any progress in understanding favoring in terms of an explanation of what one ought to do if such an explanation presupposes an antecedent grasp of the notion of favoring, or if facts about favoring are more basic than facts about having a weight in an explanation of why someone ought to do something.

Kearns and Star propose an alternative account, according to which reasons are *evidence* for ought facts. As they put it, “a fact F is a reason just in case it is evidence that one ought to  $\phi$ , for some particular  $\phi$ , where this  $\phi$  is either a belief or an action” (Kearns and Star 2009, 217). Evidence that one ought to  $\phi$ , in turn, is that which *raises the probability* of the hypothesis that one ought to  $\phi$ . Kearns and Star argue that epistemic and practical reasons are “of a kind”, and epistemic reasons are plausibly evidence that we ought to believe something. They thus “extend a plausible analysis of epistemic reasons to all normative reasons” (222) and challenge the opponent to provide an alternative common feature, other than the uninformative notion of favoring.

Kearns and Star's insight is that as favorers, reasons weigh in favor of doing something, and consequently in standard cases raise the probability that one ought to do it. However, it seems that not everything that raises (or lowers) the probability that one ought to do something is a reason to do it (or against it). Suppose Arya is in trouble and I am able to help her. The first fact plausibly favors my helping her, but how about the second? It doesn't seem so – by itself, being able to help is no reason to help. Yet my being able to help does raise the probability that I ought to do so, since it wouldn't be the case that I ought to help if I couldn't do so. In Dancy's (2004a) language, my ability to help is an *enabler*: it 'turns on' the reason provided by Arya's need for help. Or take a case from Kearns and Star themselves:

Stephen is in pain and Daniel has an unlimited supply of easily accessible pain-killers that he could give to Stephen. Daniel knows this and thus has a reason to give Stephen the pain-killers. However, Stephen is indifferent to his pain. Daniel knows this. (Kearns and Star 2013)

In this case, the fact that Stephen is indifferent to his pain lowers the probability that Daniel ought to give him painkillers, and thus amounts to evidence against it being the case that Daniel ought to do so. But it plausibly doesn't *favor* not giving Stephen painkillers. The fact that Stephen is indifferent to his pain is an *attenuator* or *disabler*: it weakens or undermines the reason-giving force of the fact that Stephen is in pain (see also Brunero 2009). Although Stephen's indifference is evidence that it is not the case that Daniel ought to give the painkillers, it doesn't *weigh against* the reason to give the them – indeed, if it is a disabler, there is, in this case, no reason to give painkillers to begin with. Relatedly, when Daniel deliberates about what to do, it is not appropriate for him to weigh the indifference against the pain (contrary to what Kearns and Star 2013 claim). Rather, he should think that in this case, he has less reason to give the painkillers than he would in ordinary contexts. As Broome emphasizes, reasons for  $\phi$ -ing are weighed against reasons against  $\phi$ -ing, while evidence that one ought to  $\phi$  is weighed against evidence that it is not the case that one

ought to  $\phi$  (rather than evidence that one ought not  $\phi$ ) (Broome 2008, 102-3). Since some evidence that one ought to  $\phi$  doesn't favor  $\phi$ -ing and plays a different deliberative role from reasons to  $\phi$ , it seems that not everything that is evidence for an ought is a reason.

On the other hand, it also seems that not all favorers increase the probability that one ought to  $\phi$ , and thus constitute evidence for it. To see this, consider Kearns and Star's proposal for making sense of the *strength* of a reason in terms of strength of evidence: "The strength of a reason to  $\phi$ , R, depends on the degree to which R increases the probability that one ought to  $\phi$ . The more probable it is that one ought to  $\phi$  given R, the stronger reason to  $\phi$  R is." (2009, 232) This account fails if there are reasons  $r$  such that the probability that one ought to  $\phi$  is higher in the absence of  $r$  than in its presence. Dancy (2004a) has argued that there are indeed such cases. He observes that I have a reason to lend someone a hand because he's my friend, but I might have even more reason to lend a hand if the person wasn't my friend, because it would be more generous and admirable. So even if  $r$  is a reason for me to  $\phi$ , I may have stronger reason to  $\phi$  in the absence of  $r$  than given  $r$ . So it may be that the probability that one ought to  $\phi$ , given  $r$ , is actually lower than the probability that one ought to  $\phi$ , given not- $r$  – even though  $r$  is a genuine reason to  $\phi$ . So some reasons seem not to constitute evidence for (raise the probability of) an ought.

So while Kearns and Star are successful at raising serious doubt about Broome's claim that reasons are to be understood in terms of a weighing explanation of oughts, there are also significant worries about Kearns and Star's claim that reasons are evidence.<sup>2</sup> Although none of the objections can be regarded as decisive, these problems with leading reductivist accounts motivate searching for an alternative approach.

---

<sup>2</sup> In their 2013 paper, Kearns and Star offer responses to related criticisms by Broome and Brunero. I lack the space for a proper discussion, but I believe that even if their responses to earlier criticisms are successful, they are not effective against the type of cases I discuss.

## 2. The View

Critics of reason primitivism propose to characterize the reason-role in terms of other normative concepts or properties without circularity. The brief examination in the previous section of some of the challenges facing the leading reductive proposals highlights the desiderata that a satisfactory account must meet. We must, first of all, be sure to identify reasons facts in a way that does not already presuppose a grasp of what favoring is. We must also take care to make sense of the distinction between favorers and other sort of facts that also make a difference to what one ought to do, as well as the relative weight of reasons. More generally, identifying reasons facts with other normative facts should illuminate the nature of reasons, to help us understand what it is for one thing to favor another in the first place.

My new proposal is inspired by what Gibbard (1990) says about *thinking* about reasons. For Gibbard, reason thoughts consist in an overall attitude toward giving weight to a consideration. According to him,

To say that R is some reason for S to  $\phi$  in C is to express acceptance of a system of norms that directs us to award some weight to R in deciding whether to  $\phi$  in C.  
(Gibbard 1990, 163)

I'm not interested in the expressivist aspect of this suggestion, but only in the link between having some reason and awarding some weight in deciding. It is a natural extension of Gibbard's suggestion to say that R *is* a reason for S to  $\phi$  in C iff S *ought* to award some weight to R in deciding whether to  $\phi$  in C – after all, what S ought to do is what the *correct* system of norms tells her to do.

The suggestion on the table is roughly that R is a reason for S to  $\phi$  iff S ought to place some positive weight on R in deciding or deliberating whether to  $\phi$ . Note that it involves no commitment to what if anything makes it the case that S ought to do so. As such, it's compatible with any view about the source of oughts – they could be primitive non-

natural facts, or reducible to some other kind of fact. But what exactly does awarding positive weight to something amount to? If it is just a matter of regarding something as a consideration favoring a certain course of action, we have made little headway. As we already saw in the discussion of Broome, if we assume a grasp of the favoring relation, we presuppose precisely what most needs explaining.

I believe the surest way to avoid circularity is to frame the view in strictly causal terms. Placing or awarding weight on a proposition should be understood simply as a matter of awareness of the proposition having a causal influence on the subject. I will also drop the mention of deciding or deliberating to increase the generality of the analysis. I'll begin with *normative reasons*, reasons that potentially add up to an ought, can justify an action or belief, and are linked to fitness of blame and praise towards the agent (see next section). Here's what I'll call the Mechanism Formulation of my view of normative reasons:

Pro Tanto Normative Reason (Mechanism Formula) (PTNR):  $r$  is a pro tanto normative reason for A to  $\phi$  if and only if (and in virtue of the fact that) A's (actual or counterfactual) awareness of  $r$  *ought to* have causal influence in the direction of  $\phi$ -ing on the psychological mechanisms responsible for A's  $\phi$ -ing or not  $\phi$ -ing.

The analysis is intended to be general so that it applies to epistemic as well as practical reasons, although I will focus on the latter here. According to it, that it is raining is a pro tanto normative reason for Samantha to open her umbrella if and only if Samantha's awareness of the fact that it is raining ought to causally influence the mechanism responsible for her potentially opening her umbrella in the direction of opening it.

Let's leave it open why the fact that it is raining ought to influence Samantha and focus on the analysis. To begin with, we have a fact, a true proposition: it is raining where



Samantha is, and a potential action: opening the umbrella. This fact, by itself, won't causally influence Samantha in the relevant way. It is only if she is aware of the fact that it can make a difference to her reason-responsive mechanisms. I take it that awareness need not be conscious: we can be responsive to considerations without explicit thought. The analysis says that it doesn't matter if Samantha is actually aware of the fact. There is a reason for her to open the umbrella, if and only if she ought to open it, were she aware of it in some suitably close possible world. (I address some conditional fallacy –related worries below when discussing purported unfollowable reasons.) Some might want to restrict the reasons someone *has* to those the subject is aware of. The account can easily be modified to suit the best answer to this question. In any case, awareness is a factive notion: if it's not the case that it is raining, that it is raining can't be a reason for her. That's as it should be: something that's not the case can't be a reason.

Suppose, for simplicity, that Samantha is aware of the fact that it is raining. PTNR says that if the rain fact is a reason, the awareness ought to influence her in the direction of opening the umbrella, or more precisely influence whatever psychological mechanism is responsible for her so acting. I don't have a theory of such mechanisms in general. But it's obvious enough that there are various psychological mechanisms that take as input one's beliefs or perceptions and output intentions or attitudes. One such mechanism is deliberation or practical reasoning. Engaging it may be the distinctively human way of arriving at intentions and actions, but it's surely not the only one. There are many action-guiding and attitude-guiding mechanisms and many ways for awareness of a fact to causally influence them.

Nevertheless, deliberation is in some ways in a privileged position in the case of action, since it is the process of explicitly considering reasons for and against. I take

deliberation to be a complex causal process that takes as inputs contents of beliefs and preferences, puts them together in a more or less (decision-theoretically) rational manner under constraints provided by prior intentions and values, and yields an intention as output. Since deliberation involves so many elements, there are many ways to causally influence it in some particular direction. Varying the salience of an outcome, preferences between outcomes, attitudes to risk, or ruled-out behaviors will potentially change the output of deliberation. In the rain case, perhaps Samantha's awareness of the fact that it is raining ought to raise the salience of her preference for staying dry and the easy means of achieving this goal, thus making it more likely that she will open her umbrella. Importantly, dispositional changes are also ways of weakly influencing the deliberative mechanism. You might never actually need to choose between voting Putin or Medvedev as president, but nevertheless be disposed to rank voting for Medvedev higher, were the question to arise.

The way I have formulated the analysis invites a serious objection: it is *agents* who ought to do this or that, such as respond to or place weight on something, but I only talk about what *ought to be* the case – it ought to be that a certain kind of awareness has a kind of certain causal influence. How can these claims be reconciled? According to one view, if something ought to be the case, someone ought to see to it or ensure that it is the case. But how do we see to it that our psychological mechanisms respond in a certain way? 'Ought' is supposed to imply 'can', but do we have that kind of ability?

There's a good reason to think that we do, in the relevant sense. Let's start with something that seems to be unproblematic in this respect: sometimes I ought to give weight to something in deliberation. But what exactly is it that I *do* when I give weight to something in deliberation? Well, for example, I pay attention to it, or I let it move me towards doing something. What I want to say is that there is no difference between my paying attention to

some F and that F having a particular kind of causal influence in the process of deliberation. For *me* to let F move me is just for F to (non-deviantly) causally influence a psychological mechanism that, in part, constitutes who I am, or my perspective on the world. I don't deny that we are here at the very limits of agency, the borderline between activity and passivity. On this view, there are causal chains such that their outcomes are my doings. But if we don't believe in a *sui generis* kind of agent-causation, we'll *have to* believe something like this in order to believe *we* can be causally efficacious.<sup>3</sup>

I'm going to reformulate Pro Tanto Normative Reason in explicitly agential terms to make it clearer that the ought governs the something the agent does. I'll call this the Agential Formulation of PTNR, and use PTNR' to refer to it:

Pro Tanto Normative Reason (Agential Formula) (PTNR'): *r* is a pro tanto normative reason for A to  $\phi$  if and only if (and in virtue of the fact that) A *ought to* conduct her  $\phi$ -relevant thinking in a (more or less)  $\phi$ -friendly manner, given awareness of *r*.

I emphasize that apart from leaving some qualifications tacit, PTNR' says *nothing more* than PTNR, as I understand it. ' $\phi$ -relevant thinking' is short for psychological processes that either result in  $\phi$ -ing or not  $\phi$ -ing, such as deliberation about whether to  $\phi$ . For someone to conduct her  $\phi$ -relevant thinking in a  $\phi$ -friendly manner, or equivalently to take *r* positively into account in her  $\phi$ -relevant thinking, is for her to adjust or bias her  $\phi$ -relevant thinking towards  $\phi$ -ing, one way or another, to a greater or smaller degree. Sometimes such adjusting or biasing is what we might call a basic mental action, sometimes it requires indirectly ensuring that one responds that way, such as training oneself to notice some consideration. All these phrases are just ways of talking about how causal processes in the mind ought to

---

<sup>3</sup> For a thorough compatibilist account of psychological mechanisms compatible with responsibility, see Fischer and Ravizza (1998).

run. So it adds nothing to the above to say that given her awareness of the fact that it is raining, Samantha ought to conduct her umbrella-opening relevant thinking in an umbrella-opening-friendly manner.

To ward off a misunderstanding I have occasionally encountered, let me stress that the fact that it is raining is not, on my view, a reason to think in an umbrella-opening-friendly way. It is, precisely, a reason to open the umbrella. But Samantha *complies with* her reason when she thinks in an umbrella-opening-friendly manner. If the reason is conclusive and she complies with it, she will open her umbrella, unless some external circumstance interferes.

At this point, some may have a further worry: If conducting or adjusting one's thinking is itself an action we *ought* to perform in a certain way, must there not be a *reason* for it? And doesn't that mean that reasons are after all fundamental? No. There is certainly something that makes it the case that one ought to think in a certain way. Suppose, for simplicity, that what makes it the case that Samantha ought to think in an umbrella-opening-friendly way (for short, in an U-friendly way), given that it's raining (call this fact *r*), is that Samantha's ideal advisor would want her to do so, given the rain.<sup>4</sup> (The truth or falsity of this toy explanation doesn't matter here.) Call this fact about the ideal advisor's attitudes IA. Is IA itself a reason for Samantha to think in an U-friendly way? No. IA doesn't *favor* thinking in an U-friendly way, but rather explains why *r* favors opening the umbrella (via explaining why one ought to think in an U-friendly way). IA is *not* the sort of consideration one could or should respond to, or deliberate about. It is thus not a reason in the sense under analysis, but simply part of a metaphysical explanation of why an ought obtains. In Schroeder's (2007a, ch. 2) terms, it's a background condition, an *explanatory* rather than a

---

<sup>4</sup> This type of view is famously defended by Michael Smith (1994).

*normative* reason. At least some oughts governing an agent are explained otherwise than in terms of the balance of favorers, which allows them in turn to explain favoring.

So far, I have focused on normative reasons. PTNR captures a central way in which a fact can favor an attitude. But there is another flavor of favoring. It is highlighted by Jonathan Dancy's discussion of what he calls enticing reasons:

Enticing reasons are to do with what would be fun, amusing, attractive, exciting, pleasant, and so on. ... But (as I understand the matter) they never take us to an ought; it is not true of an enticing reason that if one has one of them and no reason of any other sort, one ought to do what the reason entices one to do. (Dancy 2004a, 21)

Such reasons are not felicitously understood as belonging to the same category as normative reasons. Failing to respond to them isn't subject to criticism in the same way, as I'll discuss below. They are better treated as exemplifying a different kind of favoring. I will call them *evaluative reasons*. Here is my proposal:

Pro Tanto Evaluative Reason (Mechanism Formula) (PTER):  $r$  is a pro tanto evaluative reason for A to  $\phi$  if and only if (and in virtue of the fact that) it is *desirable* that awareness of  $r$  has causal influence in the direction of  $\phi$ -ing on the psychological mechanism(s) responsible for A's  $\phi$ -ing or not  $\phi$ -ing.

Again, the parallel agential formula PTER' says that  $r$  is a pro tanto evaluative reason for to A to  $\phi$  if and only if it is desirable for A to conduct her  $\phi$ -relevant thinking in a (more or less)  $\phi$ -friendly manner, given awareness of  $r$ . The idea is that when it is desirable – whether prudentially, morally, or aesthetically – for someone to be responsive to some consideration by  $\phi$ -friendly thought, the consideration throws one kind of positive light on  $\phi$ -ing.

Nevertheless, someone who fails to conduct her thinking in a desirable way is not thereby subject to blame or criticism, so evaluative reasons are not normative in the narrow sense I employ here. Enticing reasons are not the only sort of evaluative reason. Supererogatory

reasons, such as the reason I have to give most of what I earn to the poor, behave similarly. Perhaps evaluative reasons are enticing when it is *prudentially* desirable to be influenced by them and supererogatory when it is *morally* desirable to be influenced by them.

The relation between normative and evaluative reasons raises big questions I will not be able to address here. Suffice it to say that there will be no single answer to the question of what an agent has most reason to do when normative and evaluative reasons point in different directions. Nevertheless, this type of incommensurability is in one sense benign: since evaluative reasons do not bear on what the agent *ought* to do, it will always be clear that the agent ought to do what she has most normative reason to do. Evaluative reasons play their role in guiding action when the agent has more than one permissible option open.

### **3. The Rationale**

I have proposed that reasons facts are facts about how we ought to be or it would be desirable for us to be causally influenced, which are facts about how our thinking ought to be conducted or how it would be desirable to conduct it, given awareness of the fact that is a reason. I claim that this gives us an informative account of the favoring relation. One might nevertheless wonder why we should bother – why should we care whether pro tanto normative notions are reducible to all-things-considered notions? The answer is that the reduction helps explain the sense in which reasons are *normative* or *evaluative*.

Begin with normativity. To some ears, the very question may sound strange. Recall the quote from Raz: “The normativity of all that is normative consists in the way it is, or provides, or is otherwise related to reasons.” (1999, 67) So isn’t it obvious that reasons are normative? No, it isn’t. Think about what it is for something to be normative. This is not a simple question, and there is no consensus on what the normativity of the normative consists in. But there is good reason to think that it comes down to the fittingness or appropriateness

of certain reactive attitudes and other forms of crediting or discrediting agents. One way to put this is that normative facts are *action-guiding* because they are *reaction-guiding*. When we fail to do something we are supposed to do, we are liable to blame or criticism, by ourselves as well as others. That is what the ‘normative pressure’ (Kolodny 2005) of reasons consists in. Normative reasons have *authority* for us, and this derives, at least in large part, from our being a fitting target of recrimination by our better selves and those we respect if we fail to take them into account. The tie between reasons and justification is also related to licensing reactions. When something justifies an action or belief, it shows that negative reactions towards the agent on account of the action or belief are out of place, even in the absence of an excuse. Maybe some kind of positive reaction, such as trust, is appropriate.

So, let’s say that for what it is for something to be normative is for it to provide an authoritative standard that we are accountable for living up to, that is, acting in accordance with. There are, to be sure, excuses and exemptions. But roughly, for you to be accountable for something, it must be something you should have done and could have done but didn’t do. If it is the case that you ought to  $\phi$ , it is clear that what you are accountable for is  $\phi$ -ing. It is fitting to respond negatively to you for not  $\phi$ -ing, unless you have an excuse. It is clear what it takes to comply with an ought. But what does it take to comply with a pro tanto reason – what does the reason provide an authoritative standard for? Suppose you are in the shop, and you can buy either margarine (A) or butter (B), but not both. Let’s say the weight of your normative reason for A-ing at  $t$ ,  $r_A$ , is 5 and the weight of the normative reason for B-ing at  $t$ ,  $r_B$ , is 8. *Qua* a normative reason, your reason to buy margarine is normative for you. But how and why, when you clearly should not buy margarine? That’s the question for the reasons primitivist. What kind of standard can  $r_A$  provide for you to act in accordance with, when it’s *not* the case that you should A? It doesn’t ‘command’ or direct you to A, as Dancy nicely puts it:

An overall 'ought' (that is, the answer to the question what we ought to do, all things considered) is indeed a directive ... It expresses a command, maybe, it has a compelling quality, it has 'obligatory force' in some sense or other, it feels like an order and directs us to do this rather than that. The question then is whether contributory reasons do the same as the overall 'ought' does. And the answer is surely that they don't. If an overall 'ought' commands, it cannot be that contributory reasons command as well; there is no such thing as a pro tanto command. (Dancy 2004b, 105)

So: if reasons don't command or direct us, what are we accountable for when we ignore them? If there's no command, how can I 'obey' or 'disobey' a reason or be under normative pressure? If I can't disobey, how can reasons be normative?

My view offers a response. There is after all something you should and could do (all things considered): you ought to conduct your A-ing relevant thinking in a somewhat A-friendly manner. We might say it should be A-friendly to degree 5 in my example. I have discussed above what this might amount to – perhaps you should give buying margarine some deliberative attention, either in this world or in some nearby possible world, instead of completely ignoring it. That's what you are accountable for, given  $r_A$ , even though you are not accountable for failing to A, since  $r_B$  is weightier (assuming A-ing and B-ing are mutually exclusive). In fact, it is to your credit that you don't A. In this way, when reasons are analyzed in terms of a special sort of thing you ought to do – taking certain considerations into account – we can see how they are linked to credit and accountability. That's what it is for reasons to be normative.

So, if I'm not causally influenced by awareness of a normative reason-fact in acting, a negative reaction towards me is fitting in the absence of an excuse or exemption. Is this plausible? Let's bear in mind that there is a broad spectrum of negative reactions that constitute holding accountable. At the one end, there are various blame-constituting emotions from indignation and outrage to resentment. At the other, there are subtle discrediting reactions from avoidance to thinking less highly of the person. It is a plausible



principle that the stronger the causal influence awareness of a fact ought to have, the stronger the negative reaction its absence licenses or requires. Second, these responses need not be mandatory. It may be that they are simply permissible in virtue of failure to respond to the reason. This is important, since we do plausibly have lots of weak reasons to do all kinds of things at any given time (more on this below). These reasons can be normative and hence have implications for accountability even if we are not required to adjust our attitudes toward each other and ourselves in response to each and every one of them.

So far, I've argued that linking normative reasons to oughts governing thinking and hence to fittingness of holding accountable helps us see the sense in which they are normative for us. Similarly, linking evaluative reasons to desirability of thinking in a certain way and hence to fittingness of desiring to be a certain kind of thinker helps us see the sense in which they favor action. The fact that smoking a pipe at Les Deux Magots would be cool is an evaluative reason for me to do so. According to PTER, this reason's fact just is the fact it would be (prudentially) fitting for me to want to be the sort of person who thinks in a Les-Deux-Magots-pipe-smoking friendly way. If I do think in such a fashion, it will also be fitting for others relevantly like me to want to be the sort of person I am – to desire to emulate me in this respect. There is a kind of agential value realized in responding to the consideration in this way, and that's what makes the consideration an evaluative reason.

Though failure to respond to an evaluative reason does not merit a negative reactive attitude, there is an analogue. If I am the sort of person who shuns playing a computer game in spite of knowing that I would enjoy it, it is fitting for others to reduce the strength of their desire to be like me (if possible), and their desire to recommend emulating me to others. The way I respond to evaluative reasons can discredit or credit me in this sense. The fact that a certain kind of negative or positive response is licensed in virtue of my conducting my

thinking in a certain way in response to the presence of a consideration makes intelligible why evaluative reasons belong in the same broad category of favorers as normative reasons do.

These links between responding to reasons and fittingness of negative or positive attitudes should caution us against what I will call the *positive reason existential fallacy*: conflating the fact that taking an option would result in something good for an agent with there being a reason to take that option. (I choose the label by analogy with Mark Schroeder's opposite thesis of negative reason existential fallacy, according to which there is a pragmatic explanation of why we underestimate the number of weak reasons.<sup>5</sup>) This is a fallacy, because normative reasons are linked to accountability, while the mere fact that something would be good is not. That there is something desirable about an action or its consequences does not automatically provide me with a standard I can be held to. In the case of evaluative reasons, the point is a little more subtle: what would be good for me and how it would be good for me to think given my awareness of a fact are two different things. I need be no worse as an agent for ignoring some good-making feature.

#### **4. The Challenges**

I've proposed that for something to be a normative reason for someone – for a fact to favor an action or belief – is for it to be the case that awareness of it ought to causally influence the subject's actions or beliefs. Evaluative reasons are facts it is desirable for someone to take positively into account in their action-relevant thinking. This is to reduce facts about normative reasons to facts about how people ought to be, which are facts about what they ought to do in terms of taking something into account, and facts about evaluative reasons to

---

<sup>5</sup> See e.g. Schroeder's unpublished paper 'The Negative Reason Existential Fallacy' ([http://www-bcf.usc.edu/~maschroe/research/Schroeder\\_Negative\\_Reason\\_Existential\\_Fallacy.pdf](http://www-bcf.usc.edu/~maschroe/research/Schroeder_Negative_Reason_Existential_Fallacy.pdf)), accessed June 3, 2013.

facts about how it would be desirable for them to think, given their awareness of the fact that is a reason. Accounts of this general type face formidable objections, however. It is not obvious that they can make sense of some putative key features of reasons. I will discuss six challenges in this section.

#### 4.1 *Pro tanto* nature

The first challenge is explaining the *pro tanto* nature of reasons – the fact that we can have reason to do something even if we ought not, all things considered, do it. Unsurprisingly, PTNR has no problem with explaining this. It is possible that a psychological mechanism, such as deliberation, ought to be influenced by S's awareness of  $r$  toward  $\phi$ -ing, even if S ought not to  $\phi$  all things considered. This is because a mechanism can be influenced towards  $\phi$ -ing while nevertheless being more strongly influenced by countervailing considerations. The same goes for PTER. So this feature of reasons is no problem for the account.

#### 3.2 *Weight*

The second challenge is to make sense of the *weight* of reasons. It is a natural extension of the view to say that the *strength* or *weight* of a reason is a matter of how much causal influence awareness of a fact ought to have (or how much causal influence it is desirable for the fact to have, in the case of evaluative reasons). This allows us to make sense of comparisons of the weight of reasons. I will focus here on normative reasons. Extending the account to evaluative reasons, *mutatis mutandis*, should be straightforward.

So, the view is that the fact that Ringo is about to drown is a stronger reason for George to save Ringo than the fact that he has promised to help John is to help John, because his awareness of the former ought to be influence him more than his awareness of the latter. To put it more generally:

Weight: For  $r$  to be a stronger normative reason to  $\phi$  than  $r'$  is to not  $\phi$  is for it to be the case that awareness of  $r$  ought to influence the mechanism responsible for  $\phi$ -ing or not in the direction of  $\phi$ -ing to a greater degree than awareness of  $r'$  ought to influence the mechanism in the direction of not  $\phi$ -ing.

One attractive thing about this account of the weight of normative reasons is that a natural relationship between what one has most normative reason to do and what one overall ought to do falls out of it. For the balance of reasons to favor  $\phi$ -ing is for it to be the case that the resultant, as it were, of all the causal influences that ought to make a difference to  $S$ 's  $\phi$ -ing or not  $\phi$ -ing points toward  $\phi$ -ing. If that's the case, then if  $S$  is the way she ought to be, she will  $\phi$ , if she can. That is to say that  $S$  overall ought to  $\phi$  in the sense of having conclusive reason for it.

### *3.3 No motivational or deliberative isomorphism*

If the preceding challenges are relatively easy for my type of account, the challenge from lack of isomorphism is targeted precisely at views like PTNR and PTER. We have lots of reasons for doing this and that. And as Dancy puts it, "It is not as if for every reason one ought to have a little bit of motivation [...] am I somehow in the wrong if I am not motivated at all by a reason which I am well aware is defeated in the present case?" (Dancy 2004a, 21) I agree that it is important to accommodate this. As a softening up point, let me note that the supposed datum that there are millions of tiny normative reasons for each of us is open to question. We must bear in mind the danger of the positive reason existential fallacy. I believe that Schroeder, for example, falls prey to it when he talks about my reason to eat my car – namely, the fact that it contains the recommended daily allowance of iron (2007b, 123). It is true that this is a good-making feature of the potential action. But it provides no

standard for me to comply with, nor is responding to it a good-making feature of an agent. It is not just that it is pragmatically odd to say it is a reason, as Schroeder speculates – it simply is not a normative or evaluative consideration of the right kind to be a reason for me.

Even if we cull the herd of putative reasons, it is nevertheless plausible that many weak reasons remain. How can it be the case that one ought to be influenced by  $r$  in the direction of  $\phi$ -ing when it is not the case that one ought to be motivated to  $\phi$  or take  $r$  into consideration in settling whether to  $\phi$ ? The answer, implied by what I've already said about psychological mechanisms, is that there are many ways of causally influencing mechanisms responsible for action that do *not* all amount to motivation or deliberative consideration. Suppose I have a weak normative reason to break into a Beyoncé dance routine in the middle of a philosophy talk (it would provide some much needed levity and increase everyone's well-being a bit). In ordinary circumstances, it is implausible that I should be at all motivated to do so, or deliberate about it. But insofar as there is indeed some reason for me to do so, it does make sense that it ought to make some difference to my dispositions, and therefore to my responses in nearby possible worlds. Were the thought to occur to me, I ought not feel as much revulsion toward it as toward something I have *no* reason to do, like stabbing myself, nor should I dismiss it as quickly. Perhaps I should be disposed to give dancing some proper consideration, were the audience inexorably drifting toward sleep. In other words, I ought to think in a somewhat dancing-friendly way, while I shouldn't think in a stabbing-friendly way at all.

A related challenge concerns Williamsian 'one thought too many' cases (Williams 1981). Suppose two children are in the water screaming for help. One of them is my son. One reason to save him seems to be that otherwise I'll have no use for the new jacket I just bought for him, and I won't get my money back. It is not just that it isn't the case that awareness of this fact about the jacket ought to feature in my decision-making process. It

*ought not* so feature, and it would be *bad* if it did. There is no denying this. This kind of overwhelming reason deriving from a special relationship calls for ignoring other considerations in deciding what to do, monetary considerations certainly among them. The best way to rescue the account, then, is to deny that having no use for the new jacket otherwise *is* a reason to keep my son alive. And this is in fact an independently plausible thing to say. The relationship-derived reason, or alternatively the fact of standing in a special relationship itself, *silences* (disables) many other would-be reasons to save my son. It does not merely *outweigh* them. There is no residual normative pressure, as there would be in the case of an outweighed reason. So in context, having no use for the jacket is not a reason at all. (It may still be a minor good-making feature of the state of affairs of my son staying alive.)

### 3.4 Unfollowable Reasons

In one important respect, PTNR and (especially) PTER belong to the same family of views as Kieran Setiya's account of normative reasons. What Setiya says is that "The fact that  $p$  is a reason for  $A$  to  $\phi$  just in case  $A$  has a collection of psychological states,  $C$ , such that the disposition to be moved to  $\phi$  by  $C$ -and-the-belief-that- $p$  is a good disposition of practical thought, and  $C$  contains no false beliefs." (Setiya 2007, 12) For the objection I'll discuss, it doesn't matter how good dispositions of practical thought are to be understood. The challenge that Michael Smith raises for accounts of this type is that they make it a conceptual truth that all reasons can be followed by the agent. Smith's best counterexample to this type of account is that the fact that acting in a certain way would be spontaneous can be or provide a reason for an agent to act that way. The problem is that being moved by the awareness that something would be spontaneous appears to undermine the spontaneity of the action. A similar issue arises with surprise party cases, in which awareness of the surprise

would obviously spoil it, and thereby eliminate the reason that nevertheless seems to exist for the agent.

On this point, I'm willing to bite the bullet and say the fact that acting in a certain way would be spontaneous is not a reason for action. I suspect the tendency to think so arises from two confluences. First, suppose the spontaneous action would be fun. Then there is an (evaluative) reason to perform it, namely that it would be fun (rather than the fact that it would be spontaneous). Second, it would be good for the agent to do the spontaneous thing. Perhaps spontaneity is a good-making feature. But again, that's not the same thing as spontaneity of the action being the reason there is for the agent to perform the action. Once more, being a good-making feature is not the same as being a consideration that favoring action, since it isn't constitutively linked to accountability or evaluation as an agent. The spontaneity and surprise party cases reveal another reason to separate good-making and reasons: the agent could not *act on* the reason that something would be spontaneous, since doing so would destroy the reason. Nor can she *reason from* it. And it is a plausible, though sometimes denied, constraint on reasons that they can be acted on or at least reasoned from. When we consider the link between reasons and accountability, this is not at all surprising – it makes little sense to hold someone accountable for failing to act on a reason they could not have acted on or reasoned from.

### *3.5 Conditions and Modifiers*

In discussing Kearns and Star's view, I made use of Dancy's observation that not all features that make a difference to what we ought to do (and thus can serve as evidence one way or another) are considerations that count in favor (that is, *favorers* or reasons). There are also *conditions*: facts that either *enable* other facts to favor or *disable* them from doing so (Dancy 2004a, ch. 3). For example, the fact that I would be filthy rich if I ran faster than Usain Bolt

could be a reason for me to run faster than him. Alas, I cannot run faster than Usain Bolt. (I can't even run as fast as he does.) This inability disables the would-be favorer from being a reason for me. Further, there are *modifiers*: facts that either *intensify* or *attenuate* the strength of favorers. Dancy's example of an intensifier is being the only person around capable of helping someone in need. This makes the reason to help weightier than it would be if someone else could do it as well.

Focusing on the case of normative reasons, the challenge to PTNR is that conditions and modifiers ought to make a difference to my  $\phi$ -relevant thinking, but they are nevertheless not favorers. For example, the fact that I'm the only person around ought, perhaps, increase my motivation to help, and thus move me in the direction of helping. But it is still an intensifier, not a favorer. The answer to this challenge is simple, however. It is not that being the only person around ought to move me in the direction of helping. Rather, it is the case that given that I am the only person around, the fact that Jody needs help ought to move me more in the direction of helping than it ought to move me if I was one of many capable of doing so. Generally speaking, a modifier is a consideration that makes a difference to how  $\phi$ -friendly my  $\phi$ -relevant thinking ought to be, given the reason it is a modifier for. For example, given intensifier  $i$  alone, I should neither be moved to  $\phi$  or not- $\phi$ . (That I'm the only person around, by itself, ought to neither move me to help Jody or refrain from doing so.) Perhaps given reason  $r$  alone, I should be moved to  $\phi$  to degree 5. But given  $r$  and  $i$ , I ought to be moved to  $\phi$  to degree 8. If this is how I ought to respond to my awareness of  $i$  and  $r$ ,  $i$  is clearly not a favorer, but an intensifier.

Similarly, when some fact is a condition for a reason, it is not the case that it ought to move me either in the direction of  $\phi$ -ing or not- $\phi$ -ing. (The fact that I'm able to lift the stone Jody lies under is not itself a reason for me either to lift it or refrain from doing so.) But it makes a difference to whether or not I ought to (or it would be desirable to) conduct my



thinking in a more or less  $\phi$ -friendly way, given some  $r$ . (If it is impossible for me to do something, such as prevent the Holocaust from taking place, the fact that doing so would do a world of good is no reason for me to do it.) Conditions and modifiers thus explain differences in how I ought or how it would be desirable for me to respond to some facts. This means they play a very different role than favorers do on my account.

### 3.6 Uniformity

The final challenge is that of uniformity: Practical and theoretical reasons should be accounted for in the same way, as Kearns and Star emphasize. The natural extension of the present view of normative reasons is to say that  $E$  is an epistemic reason for  $S$  to believe that  $p$  if and only if  $S$ 's awareness of  $E$  ought to causally influence her belief-forming mechanisms in the direction of believing that  $p$  (possibly by reinforcing one's belief, if one already believes that  $p$ ). So, the large footprints in the snow are a reason for Sara to believe that a bear has walked that way if and only if her awareness of the footprints ought to causally influence her belief-forming mechanisms in the direction of believing that a bear walked by. It is tempting to cash this out in terms of it being the case that she ought to increase her level of credence that a bear went by, given the footprints. This pleasingly Bayesian view is independently plausible.

In the case of epistemic reasons, however, a well-known problem rears its ugly head: it seems possible that someone ought to increase her credence in something not because of something bearing on its truth but because of instrumental considerations (an evil demon might hurt you if you didn't believe that  $p$ ). To avoid this problem, we have to say that only cases in which the subject *epistemically ought* to increase her credence involve epistemic reasons. A similar move is needed to distinguish between different kinds of practical reasons. This, to be sure, simply shifts the burden of sorting the right and wrong kind of

reasons to the metaphysics of different kind of oughts. I undertake this effort elsewhere (Kauppinen 2014).

Here are two equivalent tentative formulations of this special case of PTNR:

Epistemic Reason: E is an epistemic reason for S to believe that  $p$  if and only if S's awareness of E epistemically ought to causally influence her belief-forming mechanisms in the direction of believing that  $p$ .

Epistemic Reason': E is an epistemic reason for S to believe that  $p$  if and only if S epistemically ought to increase her credence in  $p$ , given awareness of E.

## 5. Conclusion

I have argued that the two flavors of the favoring relation can be analyzed in terms of oughts governing how to conduct our thinking and the desirability of certain ways of conducting our thinking, given our awareness of the fact that is a reason. Facts about how we ought to think or it would be good for us to think are fit to play the reason-role, so we don't have to regard reasons facts as basic, primitive, or unanalyzable. Consequently, we do not need to embrace Reasons Primitivism.

To be sure, there are other reductivist proposals, and I have not produced a knock-down argument against them. But I have argued that understanding reasons, roughly, as facts awareness of which should causally influence us meets reasonable desiderata better than competing accounts. Indeed, we can now see how this view goes some way towards explaining the appeal of the rivals. It follows from the account of the weight of reasons given above that we can sometimes give a kind of weighing explanation for what we ought to do, all things considered: when we add up all the ways in which we ought to be causally influenced, given our awareness of the facts, the resultant will point in the direction of the action we ought to perform, all things considered. In this respect, the view is like Broome's,

but crucially, the for- $\phi$ -ing role is defined in non-circular terms: a reason  $r$  tilts the metaphorical scales in the direction of  $\phi$ -ing to degree  $d$  when we ought to think in a  $\phi$ -ing-friendly way to the degree  $d$ , given  $r$ . Second, it will be generally true that if we ought to think in a  $\phi$ -friendly way, given  $r$ , the presence of  $r$  increases the probability that we ought to  $\phi$ . Thus, in a large class of cases, reasons for  $\phi$ -ing will be evidence that we ought to  $\phi$ , just as Kearns and Star say. But since I don't identify reasons with evidence for oughts, I leave open the possibility that considerations that don't favor something can nevertheless be evidence for oughts and that considerations that do favor something may fail to raise the odds that one should act accordingly.

Finally, I have argued that my account of reasons has certain distinctive advantages. When reasons are constitutively linked to fitting attitudes toward the subject for whom they are reasons, their distinctive nature as favorers is made clear. Granted, all I have said concerns relationships among normative concepts or facts. I have not argued for any view about the nature of fittingness relations. Perhaps they are irreducible, *sui generis* normative facts. Perhaps they can be reduced to non-normative facts. My conclusion is neutral on this question. It merely says that instead of treating reasons or favoring as the primitive, those who either aim to reduce normative facts to the non-normative or find a place for the irreducibly normative should take aim at facts about fittingness rather than primitive reasons facts. Somewhat surprisingly, to understand how a fact can favor something, we have to understand how we ourselves ought to or how it would be desirable for us to function as rational agents and believers, given our awareness of the fact.<sup>6</sup>

---

<sup>6</sup> This paper benefited greatly from feedback from audiences at the Workshop on Normativity at the University of Tartu in 2012, SLACRR 2013, and BSET 2013. In particular, I would like to thank Robert Adams, John Brunero, Daniel Cohnitz, Jonathan Dancy, Ulrike Heuer, Simon Kirchin, Gerald Lang, Susanne Mantel, Francesco Orsi, Hille Paakkunainen, Andrew Reisner, Mark van Roojen, Matt Smith, Michael Smith, Pekka Väyrynen, Jonathan Way, Amna Whiston, and Eric Wiland for helpful criticisms and suggestions. In addition, I was very fortunate to receive insightful and challenging written comments on earlier drafts from Stephen Kearns, Barry Maguire, Lilian O'Brien, Daniel Star, and Jussi Suikkanen.

## References

- Broome, John (2004), 'Reasons'. In Wallace et al (eds.), 28–55.
- Broome, John (2005), 'Does Rationality Give Us Reasons?' *Philosophical Issues* 15, 321–37.
- Broome, John (2008), 'Reply to Southwood, Kearns and Star, and Cullity'. *Ethics* 119, 96–108.
- Brunero, John (2009), 'Reasons and Evidence One Ought'. *Ethics* 119, 538–45.
- Dancy, Jonathan (2004a), *Ethics Without Principles*. Oxford: Oxford University Press.
- Dancy, Jonathan (2004b), 'Enticing Reasons'. In Wallace et al (eds.), 91–118.
- Enoch, David (2011), 'On Mark Schroeder's Hypotheticalism'. *Philosophical Review* 120 (3), 423–46.
- Fischer, John and Ravizza, Mark (1998), *Responsibility and Control*. Cambridge: Cambridge University Press.
- Gibbard, Allan (1990), *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- Kauppinen, Antti (2014), 'Fittingness and Idealization'. *Ethics* 124 (3), 572–588.
- Kearns, Stephen and Star, Daniel (2008), 'Reasons: Explanations or Evidence?' *Ethics* 119 (1), 31–56.
- Kearns, Stephen and Star, Daniel (2009), 'Reasons as Evidence'. *Oxford Studies in Metaethics* 4, 215–42.
- Kearns, Stephen and Star, Daniel (2013), 'Weighing Reasons'. *Journal of Moral Philosophy* 10 (1), 70–86.
- Kolodny, Niko (2005), 'Why Be Rational?' *Mind* 114, 509–563.

- Parfit, Derek (2011), *On What Matters, Volume 1*. Oxford: Oxford University Press.
- Raz, Joseph (1999), *Engaging Reason*. Oxford: Oxford University Press.
- Scanlon, Thomas (1998), *What We Owe To Each Other*. Cambridge, MA: Harvard University Press.
- Schroeder, Mark (2007a), *Slaves of the Passions*. Oxford: Oxford University Press.
- Schroeder, Mark (2007b), 'Weighting for a Plausible Humean Theory of Reasons'. *Noûs* 41 (1), 110–132.
- Setiya, Kieran (2007), *Reasons Without Rationalism*. Princeton: Princeton University Press.
- Smith, Michael (1994), *The Moral Problem*. Oxford: Blackwell.
- Smith, Michael (2009), 'Reasons With Rationalism After All'. *Analysis Reviews* 69 (3), 521–530.
- Wallace, R. Jay, Pettit, Philip, Scheffler, Samuel, and Smith, Michael (eds.) (2004), *Reason and Value. Themes from the Moral Philosophy of Joseph Raz*. Oxford: Oxford University Press.
- Williams, Bernard (1981), 'Persons, Character, and Morality'. In *Moral Luck*. Cambridge: Cambridge University Press, 1–19.