

# ACTING AND BELIEVING UNDER THE GUISE OF NORMATIVE REASONS\*

Keshav Singh

University of North Carolina at Chapel Hill

[penultimate draft, please cite published version]

## Abstract

In this paper, I defend an account of the reasons *for which* we act, believe, and so on for any  $\Phi$  such that there can be reasons for which we  $\Phi$ . Such reasons are standardly called *motivating reasons*. I argue that three dominant views of motivating reasons (psychologism, factualism and disjunctivism) all fail to capture the ordinary concept of a motivating reason. I show this by drawing out three constraints on what motivating reasons must be, and demonstrating how each view fails to satisfy at least one of these constraints. I then propose and defend my own account of motivating reasons, which I call the Guise of Normative Reasons Account. On the account I defend, motivating reasons are propositions. A proposition is the reason for which someone  $\Phi$ -s when (a) she represents that proposition as a normative reason to  $\Phi$ , and (b) her representation explains, in the right way, her  $\Phi$ -ing. As I argue, the Guise of Normative Reasons Account satisfies all three constraints on what motivating reasons must be, and weathers several objections that might be leveled against propositionalist views.

## 1. Introduction

This paper is about motivating reasons. Motivating reasons are the reasons *for which* we act, believe, and so on for any  $\Phi$  such that there can be reasons for which we  $\Phi$ .<sup>1,2</sup> These reasons are also sometimes said to be the reasons *in light of which*, or the reasons *on the basis of which*, we act, believe and so on. In this paper, I propose and defend an account of what it is to act or believe for a reason. I

---

\* I owe thanks to Ram Neta and Alex Worsnip for providing extensive comments on several drafts of this paper throughout the writing process. I also thank Clayton Littlejohn, Julien Dutant, Geoff Sayre-McCord, Thomas Hofweber, Lindsay Brainard, Chris Howard, John Phillips, Chris Blake-Turner, Ian Cruise, Jessi Addison, David Faraci and an anonymous reviewer for helpful discussion and comments.

<sup>1</sup> A note on terminology: it's standard to use the term 'motivating reason' in this way (see for example Alvarez (2010), Dancy (2000), Parfit (1997, 2011) and Schroeder (2007)). However, it hasn't been universal. While I am calling the reasons *for which* someone  $\Phi$ -s her motivating reasons, Scanlon (1998) calls them 'operative reasons,' and Enoch (2011) calls them 'the agent's reasons' while reserving the term motivating reason for something else. In this paper, I'll use the terms 'motivating reason' and 'reason for which' interchangeably to refer to my subject matter.

<sup>2</sup> Throughout this paper, I will use the variable  $\Phi$  to stand in for anything such that there can be reasons for which we  $\Phi$ .

focus on action and belief because they paradigmatically admit of motivating reasons, but the account I offer is generalizable to any such  $\Phi$ .

The phenomenon of acting or believing for reasons is not an obscure one, but rather a central feature of agency that everyone is familiar with pre-theoretically. However, the dominant views about motivating reasons have strayed surprisingly far from this pre-theoretical sense. According to *psychologism*, motivating reasons are mental states that cause us to act, believe and so on. According to *factualism*, they are facts in light of which we act, believe and so on. And according to *disjunctivism*, they are some combination of the above, depending on the case. All of these views have implausible implications such that none of them can do justice to the ordinary concept of a motivating reason.

The account I offer in this paper *does* do justice to the ordinary concept of a motivating reason. To show this, I draw out three constraints on what motivating reasons must be. I then argue that unlike existing views, my account can satisfy all three of these constraints without any significant drawbacks. On the account I defend, motivating reasons are propositions. A proposition is the reason for which someone  $\Phi$ -s when (a) she represents that proposition as a normative reason to  $\Phi$ , and (b) her representation explains, in the right way, her  $\Phi$ -ing.

The paper will proceed as follows. In §2, I explain and defend the three constraints on what motivating reasons must be. In §3, I explain why psychologism, factualism and disjunctivism each violate at least one of these constraints. In §4, I lay out my positive account of motivating reasons, which I call the **Guise of Normative Reasons Account**. In §5, I respond to several objections to this account and defend the claim that it does full justice to the ordinary concept of a motivating reason. In §6, I conclude by briefly exploring some of the implications of my account.

## 2. What Motivating Reasons Must Be

In this section, I draw out three constraints on what motivating reasons must be. I argue for each of these constraints from pre-theoretical starting points and without assuming any particular account of motivating reasons. However, one thing I will assume throughout is that motivating reasons admit of a unified account that applies to action, belief, and any other  $\Phi$  such that it is possible to  $\Phi$  for reasons.

Before drawing out these constraints, some clarifications about *normative* reasons will also be useful. Whereas motivating reasons are the reasons for which we  $\Phi$ , normative reasons are standardly

understood as considerations that count in favor of  $\Phi$ -ing,<sup>3</sup> whether or not we actually  $\Phi$  on the basis of them (or even  $\Phi$  at all). I will use the language of normative reasons and the language of counting in favor interchangeably in this paper.

Within the ideology of normative reasons, there is also a distinction between there being a reason for some agent to  $\Phi$  and the agent's *having* a reason to  $\Phi$ . Reasons of the latter kind are sometimes referred to as possessed reasons. The relation between the two is controversial, so I will attempt to remain as neutral as possible about it.<sup>4</sup> But I will often put things, in the examples I use, in terms of agents' having or not having reasons to act, believe and so on. I will assume that it is possible for an agent to have a (normative) reason to  $\Phi$  without  $\Phi$ -ing on the basis of that reason, or even at all.

With these clarifications made, I am now in a position to present the three constraints.

### *2.1. The Good Reasons Constraint*

The first constraint is what I will call the **Good Reasons Constraint (GRC)**. It is a platitude that we don't just act and believe for reasons, but at least some of the time, we act and believe for *good* reasons. The correct account of motivating reasons must capture this platitude, so we need a constraint that ensures it will be captured.

In order to draw out this constraint, we first need to know what it is for the reason for which someone  $\Phi$ -s to be a *good* reason. I propose that we understand good reasons in the following way: someone  $\Phi$ -s for a good reason just in case she  $\Phi$ -s on the basis of a consideration that counts in favor of her  $\Phi$ -ing.

Since a normative reason to  $\Phi$  is just a consideration that counts in favor of  $\Phi$ -ing, GRC can be rendered as follows:

---

<sup>3</sup> See, for example Dancy (2000), Parfit (1997, 2011), Raz (1999) and Scanlon (1998).

<sup>4</sup> For a helpful discussion of this issue, see Comesaña and McGrath (2014).

**Good Reasons Constraint:** Motivating reasons must be the kind of thing such that it is possible for agents' motivating reasons to be normative reasons.<sup>5</sup>

This is almost as much of a platitude as the statement that we sometimes act and believe for good reasons, which is not surprising, since it's meant to capture that platitude. GRC shouldn't require much by way of further argument, because it's hard to see any grounds for denying it that are independent of trying to defend an account of motivating reasons that violates it. Thus, GRC is a genuine constraint on what motivating reasons must be.

## 2.2. *The No-good Reasons Constraint*

We sometimes act and believe for good reasons, but we also often act and believe for reasons that are not so good. They might be *bad* reasons, or they might merely fail to be good reasons.<sup>6</sup> Either way, not every reason *for which* we act or believe is a good one. This too is a platitude.

If, as I have claimed,  $\Phi$ -ing for a good reason is just  $\Phi$ -ing for a normative reason, then failing to  $\Phi$  for a good reason is naturally understood as  $\Phi$ -ing for a reason that fails to be a normative reason. Since they aren't necessarily *bad* reasons, let's call reasons that fail to be good reasons 'no-good reasons.' I propose that we understand no-good reasons as follows: someone  $\Phi$ -s for a no-good reason just in case she  $\Phi$ -s for a reason that fails to be a normative reason to  $\Phi$ .

Normative reasons are standardly understood as facts.<sup>7</sup> Thus, there seem to be two ways in which an agent's motivating reason can fail to be a normative reason to  $\Phi$ : (1) by being a fact that fails to count in favor of  $\Phi$ -ing, or (2) by failing to be a fact at all (even if it's the case that if it were a fact, it would count in favor of  $\Phi$ -ing).

---

<sup>5</sup> GRC is roughly equivalent to what Dancy (2000) calls the "Normative Constraint."

<sup>6</sup> One might think, for example, that considerations the agent mistakenly but reasonably takes to be good reasons are neither good reasons nor *bad* ones.

<sup>7</sup> Though this is the standard view of normative reasons, some argue that normative reasons are not facts, but propositions. I will simply assume the standard view that normative reasons are facts. However, the account of motivating reasons I ultimately defend is not only fully compatible with, but even congenial to, the view that normative reasons are propositions.

To make this distinction clear, consider the following two examples:

HIKE: Koah is hiking when he comes upon a mountain lion. He knows that mountain lions are dangerous, and on the basis of this, he decides (unreasonably) to fight the mountain lion. The mountain lion really is dangerous, but this fact doesn't count in favor of fighting the mountain lion, so Koah has no normative reason to do so.

MEDICATION: Alice unreasonably believes that she is having an allergic reaction, and so on the basis of this, she takes an allergy medication. If she really were having an allergic reaction, that would count in favor of taking the medication. But she isn't, so she has no such normative reason to take the medication.

In HIKE, Koah acts for a reason that is a fact but fails to count in favor of his action. In MEDICATION, Alice acts for a reason that fails to be a fact at all, even though if it were a fact it would count in favor of her action. Both are cases in which the reason for which someone acts fails to be a normative reason. Koah and Alice both act for no-good reasons. While these cases each involve actions, analogous sets of cases involving belief can be easily generated.

Each of these ways of  $\Phi$ -ing for no-good reasons is common in ordinary life. Thus, a constraint that captures our  $\Phi$ -ing for reasons that aren't good ones should accommodate both. We can render such a constraint as follows:

**No-good Reasons Constraint (NRC):** motivating reasons must be the kind of thing such that agents' motivating reasons can fail to be normative reasons, either by failing to count in favor of  $\Phi$ -ing or by failing to be facts.

NRC captures both ways in which we can fail to  $\Phi$  for good reasons. Importantly, it does not follow from the fact that someone  $\Phi$ -s for no-good reasons that she doesn't  $\Phi$  for any reason at all. Intuitively, we can have motivating reasons that fail to be normative reasons in both of the ways specified above. As with GRC, it's hard to see any grounds for denying NRC that are independent of any attempt to defend an account of motivating reasons that violates it. Thus, NRC is also a genuine constraint on what motivating reasons must be.

### 2.3. *The Supervenience Constraint*

The last of the three constraints on what motivating reasons must be is a constraint on what motivating reasons must supervene on. This is the most controversial of the three constraints, but nevertheless can be derived from pre-theoretical starting points.

One such starting point, which I hope is uncontroversial, is that an agent's motivating reason must play a motivational role (broadly understood) in her psychology. If it did not play this role in her psychology (that of motivating her, in a broad sense) it would not be her motivating reason. Thus, whatever a motivating reason is, it must be something that can play this psychological role.

I mean this starting point to be something capacious enough that it doesn't imply any particular view of what kind of thing motivating reasons are. For it isn't just mental states that play important roles in our psychology, but also the contents of our mental state(s), facts we apprehend and even bits of the external world with which we interact. So, any such kind of thing could in principle play a role in her psychology.

Someone's motivation for  $\Phi$ -ing must in this way be part of her mental life. It's hard to even imagine what it would be for something to motivate someone without interacting at all with her psychology, which suggests that it's constitutive of motivation that it is a psychological phenomenon. Of course, not all motivations are motivating reasons, because we can be motivated to  $\Phi$  by things that don't rise to the level of reasons for which we  $\Phi$ . But all motivating reasons are motivations (again, broadly understood), so motivating reasons must be part of our mental lives.

Again, this doesn't suggest that motivating reasons must be mental states, or anything else in particular, because lots of things that aren't mental states can be part of our mental lives. But it does imply that there can't be a change in what our motivating reasons are without there being some corresponding change internal to our psychology. For how could what motivates someone change while her psychology remains exactly the same? Similarly, it's difficult to imagine how two different agents who have exactly the same internal psychology could differ in their motivating reasons.

We might attempt to capture this thought with the simple claim that motivating reasons supervene on our mental states. Unfortunately, this won't quite do the trick. Following Williamson (2000), many have accepted the existence of *factive* mental states, such as knowledge. But intuitively, motivating

reasons wouldn't supervene on these factive mental states, because such mental states would not be fully internal to our psychology.

To see why this is so, we can consider familiar evil demon cases, which are commonly used to motivate the supervenience of various properties on *non-factive* mental states.<sup>8</sup> Imagine that in two possible worlds,  $w_1$  and  $w_2$ , you have the same experiences, apparent memories and intuitions, and reason in the same way to arrive at the same beliefs, actions, and so on. It seems obvious that in both worlds your motivating reasons are the same.

But now imagine that while in  $w_1$  you're getting things mostly right, in  $w_2$  you are being systematically deceived by an evil demon. If knowledge is a mental state, then your mental states differ between  $w_1$  and  $w_2$ , because in  $w_2$  you lack a great deal of knowledge that you have in  $w_1$ . But intuitively, this makes no difference to your motivating reasons; since you reason exactly the same way to the exact same beliefs and actions, it remains the case that your motivating reasons are the same in both worlds.<sup>9</sup>

So, it seems like if there are factive mental states, they aren't what motivating reasons supervene on. This is explained by the fact that factive mental states like knowledge are in some sense not wholly internal to our psychology. Similar points have been made by both Wedgwood (2002) and Schoenfield (2015), but with regard to rationality and justification, respectively.

Wedgwood argues that the rationality of belief supervenes on the "internal facts" about agents' mental states, which are the facts about our non-factive mental states and the relations between them. Similarly, Schoenfield argues that justification supervenes on "internalist conditions," which are conditions that describe agents' non-factive mental states.

---

<sup>8</sup> Such cases were popularized by Lehrer and Cohen (1983) in their presentation of what has come to be called the 'new evil demon problem' for reliabilism, but they have since been used to argue for a variety of supervenience constraints. For example, Wedgwood (2002) uses evil demon cases to argue that rationality supervenes on non-factive mental states, rather than on mental states in general. The argument I present here draws particular inspiration from Wedgwood's.

<sup>9</sup> The situation becomes a bit more complicated if externalism about mental content is true, because this would entail that the contents of non-factive mental states (and therefore also those mental states themselves) would be different in the demon world. Given that our motivating reasons are part of our mental lives, this would likely entail some difference in our motivating reasons. However, this would be a modest sort of difference. SC is compatible with the occurrence of this sort of difference, so whether or not content externalism is true does not affect the plausibility of the constraint (see footnote 12 for further discussion).

I won't take any position on the conditions for rationality or justification here. Rather, I simply want to claim that the best way to capture our pre-theoretical views about what motivating reasons supervene on is via a similar appeal to what's internal to us. Thus, we can formulate the supervenience constraint as follows:

**Supervenience Constraint (SC):** an agent's motivating reasons supervene on the internal facts about her mental states.<sup>10</sup>

I am using the term 'internal facts' just as Wedgwood does. The internal facts about agents' mental states are the facts about their non-factive mental states and the relations between them. SC, as I have rendered it, captures the intuition that there can't be a change in our motivating reasons without a corresponding change internal to our psychology.<sup>11</sup>

To be clear, SC does not entail that *nothing* about someone's motivating reasons can change without a corresponding change in the internal facts about her mental states. For example, one might think that whether or not an agent's motivating reason is a normative reason could change without any change in the internal facts. But SC is fully compatible with such extrinsic properties of our motivating reasons not supervening on the internal facts, because SC is about what the agent's motivating reason *is* intrinsically and fundamentally.

I suspect that some will balk at SC. Nevertheless, it is highly plausible, and it's hard to see what grounds there could be for denying it that are independent of any attempt to defend an account of motivating reasons (or some other philosophical thesis) that violates it. Thus, SC is a genuine constraint on what motivating reasons must be.

---

<sup>10</sup> I thank Ram Neta, Julien Dutant and an anonymous reviewer for helpful suggestions on how to formulate this constraint.

<sup>11</sup> It's important to note that accepting SC does not commit us to any objectionably strong form of internalism. It might be, for example, that when someone on Earth believes that the substance is H<sub>2</sub>O, her internal duplicate on Twin Earth believes that it is XYZ. Correspondingly, it might be that when someone on Earth's motivating reason is that the substance is H<sub>2</sub>O, her internal duplicate on Twin Earth's motivating reason is that the substance is XYZ. SC is fully compatible with such conclusions, because it only entails that our motivating reasons are *at least as* internal as our beliefs and other non-factive mental states. It does not entail that they are any *more* so, so it is fully compatible with externalist theses that entail such conclusions, such as externalism about mental content.

To sum up this section: I've explained and defended the above three constraints (GRC, NRC and SC) from pre-theoretical starting points about what it is to act, believe and so on, for reasons. None of these constraints relies on any prior account of motivating reasons for its support, and none of them by itself entails any such account. However, as I will show in §3, several of the most popular accounts of motivating reasons can't satisfy all three of these constraints, and thus should be rejected.

### 3. What Motivating Reasons Cannot Be

In this section, I argue against three dominant views about what motivating reasons are. These views are psychologism, factualism and disjunctivism. According to psychologism, motivating reasons are mental states that (in the right way) cause us to  $\Phi$ . According to factualism, they are facts the appreciation of which leads us to  $\Phi$ . And according to disjunctivism, they are some combination of the above depending on the case.<sup>12</sup>

None of these views can satisfy all three of the constraints identified in §2. So, none of these views is the correct one about what motivating reasons are. I'll consider each view in turn and show why each one violates at least one constraint on what motivating reasons must be.

#### 3.1. Psychologism

According to psychologism about motivating reasons, the reason for which someone  $\Phi$ -s is some mental state (or combination of mental states) that (in the right way) causes her to  $\Phi$ .<sup>13</sup> Different versions of psychologism say different things about which mental states motivating reasons are, and about *how* they must cause us to  $\Phi$  in order to be our motivating reasons. But all versions of psychologism hold that motivating reasons are mental states.

---

<sup>12</sup> Not all versions of these views purport to apply to *every*  $\Phi$  that admits of motivating reasons. My arguments against these views are arguments against the ability of these views to capture motivating reasons for *any* such  $\Phi$ . Therefore, I take myself to be showing that all such views are false, whether or not they are unified accounts of motivating reasons.

<sup>13</sup> For examples of psychologism about motivating reasons, see Davidson (1963), Smith (1994, 2003) and Turri (2009).

Psychologism violates GRC. According to GRC, motivating reasons must be the kind of the thing such that they can be normative reasons. Normative reasons are facts, not mental states.<sup>14</sup> So, if motivating reasons are mental states, they are not the kind of thing such that they can be normative reasons.

This is not a new point. Jonathan Dancy (2000) and others have argued extensively against psychologism on such grounds, so I won't spend much time rehashing those arguments here. But I take the fact that psychologism violates GRC to count decisively against it as an account of motivating reasons. And even those who think that's too strong should grant that violating GRC is a significant mark against psychologism, such that if another account can satisfy GRC (and the other two constraints), then we should reject psychologism in favor of that account.

### *3.2. Factualism*

According to factualism about motivating reasons, the reason for which someone  $\Phi$ -s is some fact the appreciation of which leads her to  $\Phi$ .<sup>15</sup> So, motivating reasons are not mental states but rather facts. Factualists usually claim that the way these facts motivate us is by being the contents of our mental states. Unlike psychologism, factualism satisfies GRC; because normative reasons are facts, if motivating reasons are facts then it's obvious that motivating reasons can be normative reasons. So, factualism has no problem explaining how we can  $\Phi$  for good reasons.

But factualism has a different problem. Recall that there are two ways in which an agent's motivating reason can fail to be a normative reason to  $\Phi$ : (1) by being a fact that fails to count in favor of  $\Phi$ -ing, or (2) by failing to be a fact at all (even if it's the case that if it were a fact, it would count in favor of  $\Phi$ -ing). Let's call the first 'normative error' and the second 'factual error.'

Factualism doesn't have any trouble dealing with normative error. If someone  $\Phi$ -s for a reason that fails to be a normative reason but is still a fact, this is no counterexample to factualism. The problem with factualism is that it can't make sense of factual error. In cases of factual error, the agent's putative

---

<sup>14</sup> Defenders of psychologism could, of course, claim that normative reasons are mental states rather than facts. I won't discuss this possibility here, but it does not strike me as an independently plausible view about normative reasons.

<sup>15</sup> For examples of factualism about motivating reasons, see Alvarez (2010, 2016) and Bittner (2001).

motivating reason was not a fact at all; therefore, the factualist must claim that in such cases the agent  $\Phi$ -s for no (motivating) reason at all.<sup>16</sup>

As I have shown in arguing for NRC, both normative error cases and factual error cases seem like ones in which someone fails to  $\Phi$  for a good reason, but there's still some reason for which she  $\Phi$ -s. But factualism has to deny this in cases of factual error, so factualism violates NRC.

In §2.2, I discussed two cases, HIKE and MEDICATION, which help show why NRC is a genuine constraint. In MEDICATION just as much as HIKE, there is some reason for which the agent acts that fails to be a normative reason. Thus, it's implausible to hold that in cases of factual error, the agent  $\Phi$ -ed for no reason at all. Alice clearly takes her medication for a reason, even though her (motivating) reason, which is that she's having an allergic reaction, is not a fact.

Here is an analogous set of cases with regard to belief:

PSYCHIC: Mike visits a psychic, who tells him that the lottery ticket he just bought will be the winning ticket. On the basis of this, Mike believes (unreasonably) that his lottery ticket will be the winning ticket. The psychic really did tell him that, but this doesn't count in favor of believing it, so it isn't a normative reason for him to believe it.

VACCINE: Jenny unreasonably believes that 95% of scientists agree that vaccines cause autism. On the basis of this, she believes that vaccines cause autism. If 95% of scientists really did agree that vaccines cause autism, this would count in favor of believing that vaccines cause autism. But they don't, so it's not a normative reason for Jenny to believe that.

PSYCHIC is a case of normative error, whereas VACCINE is a case of factual error. But whether the agent believes for a (motivating) reason does not seem to depend on this difference; Jenny believes

---

<sup>16</sup> It could also be argued that in cases of factual error, the agent still has a reason, but it is a different fact (say, a fact about her own mental states instead of a worldly fact). I classify this kind of view as a form of disjunctivism and discuss it accordingly in §3.3.

for a reason just as much as Mike does, even though her (motivating) reason is not a fact. It is simply implausible to claim, as the factualist does, that Jenny believes for no reason at all.

Maria Alvarez (2010, 2016) has offered a defense of factualism against such arguments from factual error. Alvarez argues that if the agent's putative reason for  $\Phi$ -ing was really no good reason at all, it is perfectly natural to say the agent  $\Phi$ -ed for no reason at all:

... it is not always clear that the right way of characterizing the kind of case just described is in terms of 'no good reason', as opposed to merely 'no reason'. For suppose that my reason for stabbing someone is, say, that I dislike his shoes. It seems wrong to say that I had *a* reason for stabbing him, only not a good reason. One might think that in this case I have no reason at all... In the example above, my reason for stabbing my victim, namely that I don't like his shoes, is not simply not a good reason to stab someone—it is rather no reason to do so. That I dislike someone's shoes is a reason; for instance, it is a reason for me not to buy the same shoes. But my dislike of the shoes is simply no reason for me to stab their wearer, even if in my aesthetic fanaticism I believe it is (Alvarez 2010, p. 143).

There are two problems with Alvarez's response. One is that she doesn't keep the concept of a motivating reason sufficiently separate from that of a normative reason. It's true that in the example above, it's natural to say that the reason for which I stab someone is no reason at all. But this locution trades on an ambiguity between two kinds of reasons. The right way to clarify such a sentence is to say that the (motivating) reason for which I stab someone is no (normative) reason at all. And that doesn't provide any support for the claim that I had no motivating reason at all.

The other problem with Alvarez's response is that it overgeneralizes. This is because it applies not only to cases of factual error but to cases of normative error as well. If Alvarez's explanation is correct, then we should think that *anytime* an agent  $\Phi$ -s for reasons that fail to be normative reasons, she fails to  $\Phi$  for any reason at all. In other words, Alvarez is committed to the implausible position that there is no such phenomenon as  $\Phi$ -ing for bad (or even no-good) reasons.<sup>17</sup>

Of course, factualists could attempt to circumscribe this explanation so that it only applies to cases of factual error. But this would be *ad hoc*. The only justification for claiming that agents  $\Phi$  for no reason in cases of factual error, but  $\Phi$  for reasons in cases of normative error, is that these are the verdicts

---

<sup>17</sup> I thank Julien Dutant for suggesting this line of argument.

factualism needs. But this justification, as we might say, would be no justification at all. Factualism violates NRC, and this counts strongly, if not decisively, against it as an account of motivating reasons.

As if things weren't bad enough for factualism, it also violates SC. Recall that according to SC, motivating reasons supervene on the internal facts about our mental states. Factualism violates SC because it entails that someone can go from having a motivating reason to having no motivating reason (and vice versa) without any change in the internal facts about her mental states. Consider the following example:

AIRPORT: Rodrigo's aunt is coming to visit him and he needs to pick her up from the airport when she arrives. Her flight is scheduled to arrive at 5 PM, and on the basis of this, Rodrigo forms the belief that she'll be ready to be picked up around 5:30. A few seconds later, the flight gets delayed and is now scheduled to arrive at 6 PM. But Rodrigo has stopped checking on the flight and doesn't see this.

Initially, it is a fact that the flight is scheduled to arrive at 5 PM, so the factualist can rightly say that Rodrigo believes for a reason (namely, the fact in question). However, a few seconds later, it's no longer a fact that the flight is scheduled to arrive at 5 PM. Since this is no longer a fact, the factualist has to say that Rodrigo no longer believes for a reason. This is implausible. Certainly, by Rodrigo's own lights, he continues to believe for a reason, and it would be quite odd to claim that he's suddenly confused about this.

SC can explain what's wrong with the above verdict. According to SC, the reason for which Rodrigo believes must supervene on the internal facts about his mental states. But according to the factualist, Rodrigo has gone from having a motivating reason to having none at all, without any corresponding change in the internal facts. In addition to violating NRC, factualism violates SC. Therefore, we should reject factualism.

### 3.3. Disjunctivism

Disjunctivism about motivating reasons is the view that agents' motivating reasons are different things depending on the kind of case.<sup>18</sup> Disjunctivists often draw the line along certain kinds of error cases, so that motivating reasons are one thing in non-error cases and another in error cases.<sup>19</sup> Some disjunctivists say that motivating reasons are facts in the non-error cases and mental states in the error cases. Others say they are worldly facts in the non-error cases and facts about mental states in the error cases.<sup>20</sup>

Here's a simple way to understand the appeal of disjunctivism: by saying different things about error cases and non-error cases, it incorporates the best of both psychologism and factualism. Disjunctivists go factualist about non-error cases, so like the factualist, they have no trouble explaining how it's possible to  $\Phi$  for good reasons. Thus, they satisfy GRC.

When it comes to error cases, disjunctivists claim either that motivating reasons are mental states or that they are facts about mental states. So, like psychologism, disjunctivism has no trouble explaining how it's possible to  $\Phi$  for reasons in cases of factual error. For example, in PSYCHIC, Mike's motivating reason would be the fact that the psychic told him his ticket would win (a worldly fact). In VACCINE, by contrast, Jenny's motivating reason would be the fact that she *believes* that 95% of scientists agree that vaccines cause autism (a fact about her mental states).

This may make disjunctivism seem very appealing. However, like factualism, disjunctivism violates SC. Recall AIRPORT, in which Rodrigo's aunt's flight gets delayed. The disjunctivist is committed to

---

<sup>18</sup> For examples of disjunctivism about motivating reasons, see Hornsby (2008) and Hyman (1999, 2015). Williamson (2000) also seems to favor some kind of disjunctivism.

<sup>19</sup> Normative error and factual error are often insufficiently distinguished in the literature on motivating reasons. In principle, disjunctivists could draw the line in all sorts of ways, but since factual error cases are the troubling ones, it would make sense for them to draw the line along cases of factual error. Sometimes they draw the line along whether the agent has *knowledge* or not, which yields similar results to drawing the line along factual error cases.

<sup>20</sup> We might call this latter version of the view 'factualist disjunctivism.' Though it holds that motivating reasons are always facts, it still merits being classified as a form of disjunctivism because it holds that they're two different kinds of fact in error cases and non-error cases. And while defenders of factualist disjunctivism don't necessarily refer to themselves as either factualists or disjunctivists, their view is closer to a standard disjunctivist view than it is to views (like Alvarez's) that are standardly thought of as factualist. But this is a terminological issue that nothing important turns on.

saying that when her flight gets delayed, Rodrigo's motivating reason for believing that his aunt will be outside the airport at 5:30 changes from the fact that her flight is scheduled to land at 5 to his *belief* that it's scheduled to land at 5 (or the fact that he believes this).

This is an implausible verdict in the same way factualism's verdict is. It certainly isn't true to the phenomenology of such cases; by Rodrigo's own lights, his reason remains the same throughout the case. And it would be quite odd to say that just because his aunt's flight got delayed, he suddenly becomes confused about the reason for which he believes. Furthermore, there's no change in the internal facts about Rodrigo's mental states, and so nothing that could explain a change in his motivating reasons.<sup>21</sup>

SC can explain why this kind of verdict is so implausible. According to SC, because Rodrigo's motivating reasons supervene on the internal facts about his mental states, there can't be a change in his motivating reason without a corresponding change in the internal facts. Disjunctivism violates SC, so we should reject disjunctivism.

The three views I've discussed in this section all have some things going for them. Psychologism makes sense of how we can  $\Phi$  for reasons that fail to be good ones, and makes sense of the fact that motivating reasons must play a particular role in our psychology. Factualism makes sense of how we can  $\Phi$  for good reasons. And disjunctivism can make sense both of  $\Phi$ -ing for good reasons and  $\Phi$ -ing for not so good ones.

But none of these views satisfies all three constraints on what motivating reasons must be, so each of them fails to capture the ordinary concept of a motivating reason. Defenders of these views might respond by claiming that these constraints simply can't all be satisfied, and that the ordinary concept (insofar as I'm right about it) is incoherent, so we must opt for a revisionary account. But this is false. In the next section, I will lay out my own account, which satisfies all three constraints and fully captures the ordinary concept of a motivating reason.

---

<sup>21</sup> Disjunctivists might attempt to appeal to content externalism to show that there *can* be changes in our motivating reasons in these sorts of transition cases. However, the kinds of changes content externalism would underwrite in such cases would not yield the extension the disjunctivist needs. So, context externalism can't save the disjunctivist here.

## 4. What Motivating Reasons Are

In the previous section, I argued that psychologism, factualism and disjunctivism all fail to capture the ordinary concept of a motivating reason. In this section, I will defend a form of *propositionalism* about motivating reasons.<sup>22</sup> As I will argue, the view on which motivating reasons are propositions can capture the ordinary concept of a motivating reason. In order to show this, it will be necessary to provide an account of how propositions get to be motivating reasons.

Thus, I will offer not just an ontology of motivating reasons, but a full account thereof. This account will make it clear how a view on which motivating reasons are propositions can capture all three constraints on what motivating reasons must be. Here is the account:

**The Guise of Normative Reasons Account:** an agent  $\Phi$ -s on the basis of some (motivating) reason R just in case: (a) she represents R as a normative reason to  $\Phi$ , and (b) her representation explains, in the right way, her  $\Phi$ -ing.<sup>23</sup>

In the remainder of this section, I'll explain what these two conditions amount to, and how the Guise of Normative Reasons Account satisfies all three of the constraints given in §2.

### 4.1. Representing as a Normative Reason

The Guise of Normative Reasons Account agrees with psychologism that an agent's motivating reason is something internal to her psychology. However, her reason is not her mental state, but rather the *content* thereof. This is a highly intuitive way of thinking about motivating reasons; when we cite the

---

<sup>22</sup> The term 'propositionalism' is used by Comesaña and McGrath (2016) to describe their account of perceptual reasons, which is similar to the account of motivating reasons I present in this paper. Importantly, though, perceptual reasons are a kind of *normative* reason for belief. As far as I know, no one has yet offered a fully developed propositionalist account of *motivating* reasons. However, Comesaña and McGrath (2014), Fantl and McGrath (2009), Scanlon (1998) and Schroeder (2007, 2008) all gesture at views of this kind.

<sup>23</sup> As its name suggests, the Guise of Normative Reasons Account has its roots in various "guise of the good" theses that have existed in various forms since at least Plato. Darwall (1983, p. 32) gestures at a guise of normative reasons sort of view, and cites Kurt Baier (1978), Don Locke (1974) and D.E. Milligan (1974) as supporting similar views.

reasons for which we  $\Phi$ , we tend to cite considerations we take to count in favor of  $\Phi$ -ing, more so than whatever mental states are constituted by our taking them so.

Considerations are so named because they're the kinds of things we consider when we're reasoning about what to do, believe, and so on. So, the best way to understand considerations in this context is as propositions that are the contents of the mental states that figure into our reasoning. On my account, the reasons for which we  $\Phi$  are propositions that we represent as counting in favor of  $\Phi$ -ing.

I'm now in a position to explain what it is to represent some consideration as a normative reason to  $\Phi$ . Because a normative reason to  $\Phi$  is just a consideration that counts in favor of  $\Phi$ -ing, to represent some consideration as such is just to represent the proposition in question as standing in the counts-in-favor relation to  $\Phi$ -ing.

I have a very thin sense of representation in mind here, where to represent A as F is just to pick out A under the concept 'F'. This can be expanded to representing relations between things; to represent A and B as standing in relation L is just to pick out A and B under the relational concept 'L'. So, to represent a proposition as counting in favor of  $\Phi$ -ing is to pick out that proposition and  $\Phi$ -ing under the relational concept 'counts-in-favor.'<sup>24</sup>

An example will be helpful. Let's say I believe that it rained last night, and the reason for which I believe that is that the streets are wet this morning. On my account, I apply the concept of the counts-in-favor relation to the proposition 'the streets are wet this morning' and the belief that it rained last night, and it is partly in virtue of this that the above proposition is the reason for which I believe.

There are three important clarifications to be made here. The first is that while motivating reasons are propositions that are the contents of representational states, they aren't necessarily the contents of *beliefs* (although they can be). This is because the representations in question can (and often do) fall

---

<sup>24</sup> This way of understanding the kind of representation in question draws inspiration from Neta (ms.).

short of belief. This gets things right, because intuitively some proposition  $p$  can be the reason for which I  $\Phi$  without my believing that  $p$ .<sup>25</sup>

The second clarification is that it isn't necessary for the agent to apply the concept of the counts-in-favor relation under that description. Someone need not be able to call it the counts-in-favor relation, the normative reason relation or anything else in particular, in order to have the capacity to  $\Phi$  for reasons. The concept 'counts-in-favor' is just the concept of a simple relation of normative support, and plausibly one every agent has, though under various descriptions.

The third clarification is that the agent's representing some proposition as a normative reason to  $\Phi$  does not entail a judgment on her part that it is all-things considered what she has most reason to  $\Phi$ , what she ought to  $\Phi$ , what is right to  $\Phi$  or what is worth  $\Phi$ -ing. It only entails that from her point of view,  $\Phi$ -ing has *something* going for it, even if that something is outweighed or defeated. Thus, my account does not preclude the existence of various kinds of akrasia, including the kind of 'inverse akrasia' discussed by Arpaly (2015).<sup>26</sup>

The importance of each of the above clarifications will become clear when I respond to some objections to my account in §5, and will, I hope, help to preempt others. But for now, I'll move on to explaining the second necessary condition on a proposition's being a motivating reason.

#### 4.2. *The Right Kind of Explanation*

Representing some proposition  $p$  as a normative reason to  $\Phi$  is necessary, but not sufficient for that proposition to be the reason for which the agent  $\Phi$ -s. For she might represent  $p$  as a normative reason to  $\Phi$  without actually  $\Phi$ -ing. Or she might represent  $p$  as a normative reason to  $\Phi$ , and  $\Phi$  only on the basis of some other consideration(s) altogether. For example:

---

<sup>25</sup> Of course, one could have a very thin conception of belief, on which these representations necessarily *do* count as beliefs. But on that conception of belief, it wouldn't be implausible to think that the representations in question just *are* beliefs.

<sup>26</sup> In standard examples of akrasia, the agent  $\Phi$ -s while judging that she ought not to  $\Phi$ , or that she has decisive reason not to  $\Phi$ . This is compatible with her judging that she has some (outweighed or defeated) reason to  $\Phi$ . In Arpaly's cases of inverse akrasia, the agent  $\Phi$ -s while *incorrectly* judging that she ought not to  $\Phi$ . Such cases are also compatible with my account.

SELFISH: Jordy is considering whether or not to volunteer at a charity. He knows that the fact that volunteering would help others counts in favor of doing it, but this doesn't motivate him to volunteer.<sup>27</sup> However, he also knows that volunteering would impress a romantic interest, and on the basis of this, he selfishly decides to volunteer.

Plausibly, if Jordy knows that the fact that volunteering would help others counts in favor of doing it, then *a fortiori* he represents the proposition that volunteering would help others as counting in favor of doing it. However, he doesn't volunteer on the basis of this consideration. Rather, the reason for which he volunteers is that it will impress a romantic interest.

This shows that merely representing some consideration as counting in favor of  $\Phi$ -ing isn't a sufficient condition for that consideration to be a motivating reason. In order for a consideration to be the reason for which someone  $\Phi$ -s, it must also be the case that this representation explains, in the right way, her  $\Phi$ -ing. The proposition that volunteering would help others isn't Jordy's motivating reason because his representing it as counting in favor of volunteering doesn't explain why he  $\Phi$ -ed at all, let alone in the right way.<sup>28</sup>

It is natural to think that when we explain an agent's action, belief, or some other such  $\Phi$  by citing the reasons for which she  $\Phi$ -ed, we render her  $\Phi$ -ing *intelligible* in a certain way.<sup>29</sup> As it happens, I think this kind of intelligibility is precisely what is ordinarily yielded when the agent's  $\Phi$ -ing is explained by her representing some proposition(s) as counting in favor of  $\Phi$ -ing. For the agent herself, the action is clearly intelligible, because there is something that counts in favor of it. But even third-personally, the action has a kind of minimal intelligibility, because we can at least see how, from the agent's own

---

<sup>27</sup> I assume here that if motivational internalism of some kind is true, it isn't true in so strong a form as to preclude that judgments of the form '*p* counts in favor of  $\Phi$ -ing' can fail to give rise to any motivation.

<sup>28</sup> Even if Jordy's representation did explain why he  $\Phi$ -ed, this wouldn't necessarily meet the second condition, because there are cases that seem obviously to involve the wrong kind of explanation. Davidsonians, and others who think that reasons explanations are causal, would likely cite deviant causal chain cases as paradigmatic cases of the wrong kind of explanation. However, I wish to remain neutral in this paper about whether or not reasons explanations are causal, while recognizing that whether or not they are, deviant causal chain cases are of the sort that should be able to be ruled out by my account.

<sup>29</sup> For examples of this sort of thought about intelligibility, see Velleman (1989), Raz (1999), Lord and Sylvan (forthcoming) and Lord (forthcoming).

perspective,  $\Phi$ -ing has something going for it. Thus, to explain, in the right way, an agent's  $\Phi$ -ing is to render it intelligible in this minimal sense.<sup>30</sup>

This point is, I think, similar to what Davidson had in mind when he wrote that because “from the agent's point of view there was, when he acted, something to be said for the action,” there is a “somewhat anemic sense” in which such an explanation justifies the action (pp. 690-691). However, I use the language of intelligibility rather than that of justification to bring out just how minimal of an achievement it is simply to  $\Phi$  for any reason at all. An action or belief that is intelligible can still be criticized as irrational or unjustified, not to mention irresponsible, foolish, delusional, and so on.<sup>31</sup>

In fact, an agent's  $\Phi$ -ing *must* be in this minimal sense intelligible to be criticized as irrational. This is because for an agent's action to be minimally intelligible is for it to be fit for further normative appraisal (as, for example, rational or irrational). And when we explain someone's  $\Phi$ -ing by citing the reason for which she  $\Phi$ -ed, this is precisely the way in which we render her it intelligible; we make it fit for appraisal as rational or irrational. In other words, citing the reason for which someone  $\Phi$ -s places her  $\Phi$ -ing in the space of reasons.

The idea that citing the agent's motivating reason provides an explanation of this kind is a familiar one that has been floated on and off in discussions of reasons explanation, though not always in terms of intelligibility.<sup>32</sup> For example, here is Dancy on the subject:

Appeal to reasons-for-which plays a specific role in the understanding, of which the idea of a feature which in the context calls for a response of a certain type is central, and any explanatory potential derives from that central idea, rather than being identical with it. An agent who acts for a reason acts in the light of a certain consideration, and

---

<sup>30</sup> An anonymous reviewer offers the following objection: imagine that someone's reason for believing that the moon is made of cheese is that Trump is president. This person may represent the proposition that Trump is president as a normative reason to believe that the moon is made of cheese. But in an ordinary sense of 'intelligible,' this doesn't render his belief intelligible, even if it explains his belief in some way. I agree that there is an ordinary sense in which his belief is unintelligible, because absent any further details, we would have difficulty making any sense of *why* he would take the proposition that Trump is president to count in favor of believing that the moon is cheese. But even so, I think there remains a more minimal sense (which I will attempt to explain further) in which the belief, though completely irrational, is still rendered intelligible by his taking “Trump is president” to count in favor of it.

<sup>31</sup> A similar point is made by Lord (forthcoming, p. 153).

<sup>32</sup> See Dray (1957), Kim (2010) and Dancy (2014).

in coming to learn what that consideration was, we understand the action from the point of view of the agent. The agent, in acting, had his eye on a certain consideration, and being told about that...we learn some part of the explanation of his so acting.<sup>33</sup>

Dancy's point can be generalized from one about action to one about any  $\Phi$  that admits of motivating reasons. Citing the reason for which the agent  $\Phi$ -s constitutes a particular explanatory achievement (what I am calling intelligibility) because in showing us what she took to count in favor of  $\Phi$ -ing, it allows us to see her  $\Phi$ -ing from her own point of view, and thus to appraise her  $\Phi$ -ing as rational or irrational.

In SELFISH, Jordy's volunteering is not rendered intelligible by his representing its helping others as counting in favor of it, even though he does represent it as such. Rather, it is rendered intelligible by his representing its impressing his romantic interest as counting in favor of it. This helps to show why the fact that it would help others is not the reason for which Jordy volunteers, despite the fact that he represents it as counting in favor of volunteering.<sup>34</sup>

These points are, I hope, somewhat commonsensical. When an agent  $\Phi$ -s for no reason at all, neither she nor anyone else can place her  $\Phi$ -ing in the space of reasons.<sup>35</sup> Neither she nor anyone else can make sense of her  $\Phi$ -ing, or assess it as rational or irrational. It is unintelligible. For someone to  $\Phi$  for some reason, then, it must be the case that citing that reason renders her  $\Phi$ -ing intelligible.

To sum up: it's necessary but not sufficient for a consideration to be the reason for which someone  $\Phi$ -s that she represents it as counting in favor of  $\Phi$ -ing. For it to be her motivating reason, that representation must also explain her  $\Phi$ -ing by rendering it intelligible.

---

<sup>33</sup> Dancy (2014), pp. 94-95.

<sup>34</sup> I have avoided providing an example where the representation does explain the agent's  $\Phi$ -ing, but nevertheless fails to render it intelligible. This is because I want to remain neutral about whether or not reasons explanations are causal, and what makes it the case that some explanation delivers intelligibility will partly depend on the answer to that question.

<sup>35</sup> Of course, if an agent  $\Phi$ -s for no reason at all, there is still *some* explanation of her  $\Phi$ -ing – some reason *why* she  $\Phi$ -ed. However, such an explanation can't play the role of showing her  $\Phi$ -ing to be intelligible; it doesn't allow us to place it in the space of reasons.

### 4.3. Constraints Revisited

Now that I've explained the basics of the of the Guise of Normative Reasons Account, I am in a position to show how it satisfies all three of the constraints presented in §2.

First, consider GRC. It says that motivating reasons must be the kind of thing such that it is possible for agents' motivating reasons to be normative reasons.

On my account, motivating reasons are propositions. Normative reasons, however, are facts. So, to satisfy GRC, motivating reasons must at least sometimes be facts. Fortunately, it's plausible to think that when propositions are true, that makes them facts. For the most part, I'll just assume that true propositions are facts. But here is one reason to think they are: as I pointed out in §4.1, considerations are best understood as propositions. So, if a normative reason is a consideration, it must be a kind of proposition. And if a normative reason is a fact, then a fact must be a kind of proposition.<sup>36</sup>

If facts are just true propositions, then it is straightforward how my account satisfies GRC. In any case of  $\Phi$ -ing for a normative reason, the consideration on the basis of which the agent  $\Phi$ -s must be a true proposition, not a false one. Therefore, it must be a fact.

Next, consider NRC. It says that motivating reasons must be the kind of thing such that the reasons for which agents  $\Phi$  can fail to be normative reasons, either by failing to count in favor of  $\Phi$ -ing or by failing to be facts.

It's fairly straightforward how my account satisfies NRC. One can  $\Phi$  on the basis of a proposition that's a fact but fails to count in favor of  $\Phi$ -ing. Or, one can  $\Phi$  on the basis of a proposition that's false and so fails to be a fact. Since these are just the two ways NRC says agents can  $\Phi$  for reasons that fail to be normative reasons, my account satisfies NRC.<sup>37</sup>

---

<sup>36</sup> It would be even simpler for my view for it to turn out that normative reasons are (fundamentally) propositions. However, since the orthodox view is that they are facts, and this paper is not about the ontology of normative reasons, I have just assumed the orthodox view that normative reasons are facts.

<sup>37</sup> As Alex Worsnip has helpfully pointed out to me, this point may help to diagnose one of the factalist's key mistakes, which is the assumption that because *some* motivating reasons must be facts, motivating reasons must *always* be facts. On my view, motivating reasons are always propositions, and sometimes these propositions are facts.

Third, consider SC. It says that motivating reasons must be the kind of thing such that they supervene on the internal facts about agents' mental states.

On my account, motivating reasons are propositions that are the contents of a particular kind of non-factive mental state. So, if someone's motivating reason changes, then so must the internal facts about her mental states. Thus, agents' motivating reasons supervene on the internal facts about their mental states, and SC is satisfied.

Of course, some things about agents' motivating reasons can change without their mental states changing, such as whether those reasons are facts or merely false propositions. However, these are mere Cambridge changes – changes in the extrinsic features of those reasons, not what the reasons fundamentally *are* – so, such changes don't violate SC.

Psychologism, factualism and disjunctivism all fail to satisfy at least one of the three constraints on what motivating reasons must be. By contrast, the version of propositionalism I've developed here, the Guise of Normative Reasons Account, satisfies all three of them. Thus, we should accept the Guise of Normative Reasons Account of motivating reasons.

## 5. Objections and Replies

In this section, I'll respond to some objections to the Guise of Normative Reasons Account. I won't be able to respond to every objection I have encountered, so I'll focus on what I take to be the most pressing worries for my account.

### 5.1. *Overintellectualization*

I've claimed that  $\Phi$ -ing for a reason involves representing the counts-in-favor relation between a proposition (the reason) and  $\Phi$ -ing. Some might worry that this overintellectualizes motivating reasons because this kind of representation is too abstract or obscure to be what is ordinarily going on when we  $\Phi$  for reasons.

I disagree. The kind of representation I have in mind isn't very abstract, and it's not at all obscure. To be able to represent the counts-in-favor relation in this way, all the agent has to do is be able to pick out a proposition and some  $\Phi$  under the concept 'counts-in-favor.'

Two clarifications from §4.1 bear repeating here. The first is that the agent's representation of a proposition as counting in favor of  $\Phi$ -ing doesn't entail the belief that it counts in favor of  $\Phi$ -ing. This *would* be an overintellectual condition on motivating reasons, but it's not what my account says.

The second is that the representation in question is a *de re* representation; the agent need not represent the counts-in-favor relation *de dicto*. In other words, as long as the concept 'counts-in-favor' is being applied, it need not be applied under that description. If the agent needed to represent the counts-in-favor relation *de dicto*, or believe the proposition being represented, this would overintellectualize motivating reasons. But my account entails neither of these conditions.

On my account, the capacity to  $\Phi$  for reasons only requires the capacity to apply a particular normative concept. Since the concept 'counts-in-favor' is just the concept of a simple relation of normative support, and plausibly one every agent has under some description or the other, it's plausible to think any agent is capable of the kind of *de re* representation required by my account.<sup>38</sup>

### 5.2. Explanatory Impotence

Another worry some might have about the Guise of Normative Reasons Account is that it can't explain the explanatory status of motivating reasons. It's become somewhat of a dogma in discussions of motivating reasons that they are a subset of so-called *explanatory* reasons, or reasons *why*. But on my account, motivating reasons are propositions, and not always true ones. The worry, then, is this: explanatory reasons must be facts, because explanations are factive.<sup>39</sup> If motivating reasons are a subset of explanatory reasons, then they also must be facts. Motivating reasons can never be false propositions, because false propositions can't explain.

I deny that motivating reasons are a subset of explanatory reasons. As I explained in §4.2, when someone  $\Phi$ -s for a reason, her motivating reason is the content of a representational state that explains,

---

<sup>38</sup> It might be further objected that in cases of small children, or perhaps some non-human animals, this will still be too intellectual, as they don't possess the concept 'counts-in-favor'. My response to this is simply to insist that, if they are really agents and can really act and believe for (motivating) reasons, then *of course* they have this concept.

<sup>39</sup> Dancy (2000) responded to a worry like this by claiming that explanations don't need to be factive. Such a response has significant drawbacks, and he has since abandoned that strategy for responding to such worries.

in the right way, her  $\Phi$ -ing. Thus, the reason *why* she  $\Phi$ -s is this representation, whereas the reason *for which* she  $\Phi$ -s is its content.

But it would be a mistake to conclude from this that motivating reasons are explanatorily impotent on my account. To think that F must be the reason why G in order to play any role in the explanation of G is to drastically oversimplify how explanation works. F can play a substantial, indeed indispensable, role in the explanation of G without being the reason *why* G, when F is a necessary component of the reason why G.

Let's apply this: if the reason *why* someone  $\Phi$ -s is that she represented some proposition as counting in favor of  $\Phi$ -ing, then that proposition is a necessary component of the explanation of her  $\Phi$ -ing, even though it's not itself the explanation of her  $\Phi$ -ing. We can't explain her  $\Phi$ -ing in the right way without citing the proposition. More precisely, we can't show her  $\Phi$ -ing to be *intelligible* without shedding light on what consideration(s) she took to count in favor of  $\Phi$ -ing. And to do this, we need to cite those considerations.

This point generalizes to all cases in which we explain things by reference to agents' mental states. In all such cases, it's not just the bare fact of the agent's being in some mental state that does the explanatory work. We need to cite the *content* of her mental state to really explain things. So, even if mental contents are abstract objects like propositions and so aren't explanatory reasons, it's clear that such things are necessary components of such explanations.

The objector may not yet be satisfied. She might further push the objection in the following way: in cases where someone  $\Phi$ -s on the basis of a fact, it seems right to say that she  $\Phi$ -s *because of* that fact. For example, if the reason for which I pick up the pace is that I'm late for work (and I actually am late for work), it seems right to say that I pick up the pace *because* I'm late for work. But doesn't my account improperly yield the verdict that instead I picked up the pace not because I'm late for work, but because I represented being late for work as counting in favor of doing so?

I agree that in the example above it's right to say that I pick up the pace because I'm late for work. Fortunately, my account can not only accommodate, but also explain, this verdict. When I pick up the pace because I'm late, it's clearly also the case that I pick up the pace because I take myself to be late. Thus, some mental state of mine *is* a reason why I pick up the pace. But, if we make the plausible

assumption that my taking myself to be late is in some sense an appreciation of my actual lateness, then the reason why I take myself to be late is that I am late. And since reasons why are transitive, this entails that the reason why I pick up the pace is that I'm late.

This doesn't just vindicate the intuition that it's right to say I pick up the pace because I'm late. It also vindicates the intuition that in non-error cases, the reason *for which* I  $\Phi$  is also the reason *why* I  $\Phi$ . If the reason for which I pick up the pace is the fact that I'm late, and the fact that I'm late is the reason why I represent that fact as counting in favor of picking up the pace, then my motivating reason is token-identical with an explanatory reason.

By contrast, my account yields the verdict that, in cases of factual error, motivating reasons are not explanatory reasons. But there's nothing strange or counterintuitive about this; it's perfectly plausible to think that one of the unfortunate things about cases of factual error is that my action isn't explained by the facts that I take to count in favor of  $\Phi$ -ing. Instead, because there are no such facts for me to be appreciating, my taking what I do to count in favor of  $\Phi$ -ing must be explained by something else.

40, 41

### 5.3. *Transparently Bad Reasons*

The third and final objection I'll consider concerns motivating reasons that one might think the agent doesn't take to be normative reasons at all. Putative cases of such reasons are ones where the someone  $\Phi$ -s for a reason that's a bad (or no-good) reason by her own lights.

---

<sup>40</sup> It might be yet further objected that this is an unacceptable conclusion because "A's reason for  $\Phi$ -ing was that  $p$ " entails "A  $\Phi$ -ed because  $p$ ." However, insofar as the former really does entail the latter, I think it only entails it in a very weak sense of 'because,' in which "A  $\Phi$ -ed because  $p$ " does not itself entail that  $p$ . Rather, as Comesaña and McGrath (2014) argue,  $p$  is a defeasible presupposition of sentences like "A  $\Phi$ -ed because  $p$ ." For example, in the right conversational context, it's fine to say something like "She went home because she had left the stove on, but when she reached home she realized she hadn't left it on after all." So, if motivating reason statements entail these very weak 'because' statements, that isn't a problem for my account. Similarly, it can be fine to say "Her reason for going home was that she left the stove on, but when she reached home she realized she hadn't left it on after all." The supposed factive entailment of motivating reason statements has been greatly overstated by factualists.

<sup>41</sup> Thus, I can also accommodate the disjunctivist conclusion about explanatory reasons defended by Hornsby (2008) and others. On my view, there's no problem with such a conclusion, because unlike *motivating reasons*, explanatory reasons aren't plausibly construed as supervening on the internal facts about our mental states. Dancy (2008) similarly suggests in response to Hornsby that a difference in the explanation of someone's action between error and non-error cases doesn't suggest a difference in her motivating reason between those cases.

Let's consider a few such cases:

MALEVOLENT: The reason for which Caligula tortures his victims is that it will inflict pain on them, yet Caligula is fully aware that this is a bad reason – it does not count in favor of torturing them.

I don't find this kind of case very convincing, because it relies for its intuitive plausibility on a conflation between distinctively *moral* reasons and normative reasons in general. Caligula could be fully aware that his motivating reason is a *morally* bad one in the sense that it's no moral reason to torture, while still representing it as a normative reason of some other kind. In fact, it's quite implausible to imagine that Caligula really doesn't see *anything at all* as counting in favor of his action, and yet also really does it for a *reason* as opposed to merely being pushed around by his psychotic impulses.

IRRATIONAL: Luke goes downstairs for the umpteenth time to check that his front door is locked. His (motivating) reason for doing this is that he might have forgotten to actually lock the door the last time he went downstairs to check. But he's fully aware that this is a bad reason and he's being irrational; it doesn't really count in favor of going downstairs.

This case is slightly more plausible than the Caligula case, because it doesn't rely on any kind of conflation between kinds of normative reasons. Ultimately, however, it's similarly implausible to imagine that Luke really doesn't see *anything at all* as counting in favor of his action, and yet also really does it for a reason, as opposed to being pushed around by a neurotic compulsion.

Luke might not *believe* that the miniscule chance that he forgot to lock the door counts in favor of going downstairs. But it's plausible to think that if this is really his motivating reason, he must on some level represent it as a normative reason to go downstairs (even if an outweighed or defeated one). Otherwise, his action would be unintelligible. If he really acts for a reason, it can't be the case that his action is unintelligible. So, either he represents it as a normative reason, or he doesn't really act for that reason. Furthermore, even if he doesn't really act for that reason, it's still possible that his thinking there's a chance he forgot to lock the door explains his action; after all, compulsions can explain actions.

Finally, let's consider an example involving belief:

DREAM: The reason for which Leia believes that God exists is that she had a dream that God spoke to her. But she doesn't think this is really any evidence that God exists, and is fully convinced that it's a bad reason to believe this – one that doesn't at all count in favor of believing it.

In cases like DREAM, we might say things like “I know this isn't a good reason to believe that  $p$ , but I just can't help believing it.” Just like in cases involving action, there are two ways to understand what's going on in such a situation. It might be that Leia doesn't really believe for any (motivating) reason at all, but rather her belief is foisted upon her by the force and vivacity of her dream-experience. Of course, if this is the case, her dream-experience could still explain her belief in some other way. On another way of understanding Leia's situation, despite knowing full well that her having this dream-experience doesn't count in favor of believing that God exists, she nevertheless represents it on some level as a normative reason (again, even if an outweighed or defeated one). One of these interpretations must be correct, and neither presents a counterexample to my account.

I suspect that all putative counterexamples of this form have the same problem. In all of them, it will be plausible to either interpret the agent as on *some level* representing the consideration as a normative reason to  $\Phi$  (even if an outweighed or defeated one), or to interpret her as not really  $\Phi$ -ing on the basis of that reason after all. So, I don't think any example of this form is a true counterexample to my account.

## 6. Concluding Remarks

In this paper, I've identified three constraints on what motivating reasons must be. According to these constraints, motivating reasons must be the kind of thing such that (a) they can be normative reasons, (b) they can fail to be normative reasons in both normative error cases and factual error cases, and (c) they supervene on the internal facts about our mental states.

I've then shown that three dominant views about motivating reasons (psychologism, factualism, and disjunctivism) can't satisfy these constraints. Psychologism can't make sense of how we can act and believe for normative reasons. Factualism can't make sense of how we can act and believe for any reasons at all in cases of factual error. And neither factualism nor disjunctivism can make sense of the connection between motivating reasons and our mental lives.

Finally, I've presented my positive account, the Guise of Normative Reasons Account. According to this account, the reasons for which someone acts or believes are propositions that she represents as counting in favor of acting or believing, where this representation explains her action or belief by rendering it intelligible.

The Guise of Normative Reasons Account satisfies all three constraints, and weathers what I take to be the most pressing objections to it. It captures the ordinary concept of a motivating reason in a manner applicable to any  $\Phi$  to which that concept can be applied, including but not limited to action and belief.

But that's not all. I'll finish by briefly gesturing at how the Guise of Normative Reasons Account is highly theoretically fruitful. It says that acting or believing for a reason involves representing some consideration as a normative reason. Such representations, like all representations, have constitutive correctness and success conditions: an agent correctly represents some consideration as a normative reason when that consideration is actually a normative reason. And she successfully represents some consideration as a normative reason when she is creditworthy for the correctness of her representation.

While the details of these constitutive conditions will have to be explored in future work, they provide the foundations of a powerful constitutivist argument for why achievements like knowledge and creditworthy action have the normative status they do. If it's constitutive of acting and believing for reasons that we represent these (motivating) reasons as normative reasons, then acting and believing for reasons have constitutive correctness and success conditions. Thus, there's an important sense in which acting and believing for reasons constitutively aim at acting and believing for *good* reasons. And if achievements like knowledge and creditworthy action can be understood in terms of good reasons, then the constitutive aims of acting and believing for reasons can explain *why* they are achievements. This would be a substantial theoretical accomplishment and a considerable victory for constitutivists.

I've only given a rough sketch of the implications of my account here, but even this rough sketch is highly promising for our theorizing about achievement in action and belief. This is all the more reason to think that to act and believe for reasons is to act and believe under the guise of normative reasons.

## References

- Alvarez, Maria. 2010. *Kinds of Reasons. An Essay in the Philosophy of Action*. Oxford: Oxford University Press.
- . 2016. "Reasons for action, acting for reasons, and rationality." *Synthese*:1-18.
- Arpaly, Nomy. 2015. "Huckleberry Finn Revisited: Inverse Akrasia and Moral Ignorance." In Randolph Clarke, Michael McKenna, and Angela M. Smith (eds.), *The Nature of Moral Responsibility: New Essays*. Oxford University Press.
- Baier, Kurt. 1978. The Social Source of Reason. *Proceedings and Addresses of the American Philosophical Association* 51 (6):707 - 733.
- Bittner, Rüdiger. 2001. *Doing Things for Reasons*. Oxford: Oxford University Press.
- Comesaña, Juan, and Matthew McGrath. 2014. "Having False Reasons." In Littlejohn, Clayton, and John Turri (eds.), *Epistemic Norms: New Essays on Action, Belief, and Assertion*. Oxford University Press.
- . 2016. Perceptual reasons. *Philosophical Studies* 173 (4):991-1006.
- Dancy, Jonathan. 2000. *Practical Reality*. Oxford: Oxford University Press (rev. ed. 2002).
- . 2008. "On how to act: disjunctively." In Adrian Haddock & Fiona Macpherson (eds.), *Disjunctivism: Perception, Action, Knowledge*. Oxford University Press. pp. 262--282.
- . 2014. "On Knowing One's Reason." In Littlejohn, Clayton, and John Turri (eds.), *Epistemic Norms: New Essays on Action, Belief, and Assertion*. Oxford University Press.
- Darwall, Stephen L. 1983. *Impartial Reason*. Cornell University Press
- Davidson, Donald. 1963. 'Actions, Reasons, and Causes'. Reprinted in: *Essays on Actions and Events*, 2nd ed., 3–19. Oxford: Clarendon Press (2001).
- Dray, William H. 1957. *Laws and Explanation in History*. Greenwood Press.

- Enoch, David. 2011. *Taking Morality Seriously: A Defense of Robust Realism*. Oxford: Oxford University Press.
- Fantl, Jeremy and Matthew McGrath. 2009. *Knowledge in an Uncertain World*. Oxford University Press.
- Hornsby, Jennifer. 2008. "A disjunctive conception of acting for reasons." In Adrian Haddock & Fiona Macpherson (eds.), *Disjunctivism: Perception, Action, Knowledge*. Oxford University Press.
- Hyman, John. 1999. "How knowledge works." *Philosophical Quarterly* 50 (197):433-451.
- . 2015. *Action, Knowledge, & Will*. Oxford University Press UK.
- Kim, Jaegwon. 2010. *Essays in the Metaphysics of Mind*. Oxford University Press.
- Lehrer, Keith & Cohen, Stewart. 1983. Justification, truth, and coherence. *Synthese* 55 (2):191-207.
- Locke, Don. 1974. Reasons, Wants, and Causes. *American Philosophical Quarterly* 11 (3):169 - 179.
- Lord, Errol. Forthcoming. *The Importance of Being Rational*. Oxford University Press.
- Lord, Errol and Kurt Sylvan. Forthcoming. "Believing for Normative Reasons: Prime, Not Composite." In Bondy, P. and Carter, A. (eds.), *Well-Founded Belief: New Essays on the Basing Relation*.
- Milligan, D. E. 1974. Reasons as explanations. *Mind* 83 (330):180-193.
- Neta, Ram. Unpublished manuscript. "The Basing Relation."
- Parfit, Derek. 1997. "Reasons and motivation." *Aristotelian Society Supplementary Volume* 71 (1):99–130.
- . 2011. *On What Matters*. Vol. 1. Oxford: Oxford University Press.
- Raz, Joseph. 1999. *Engaging Reason: On the Theory of Value and Action*. Oxford University Press.
- Schoenfield, Miriam. 2015. Internalism without Luminosity. *Philosophical Issues* 25 (1):252-272.
- Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.

———. 2008. 'Having Reasons'. *Philosophical Studies* 139 (1): 57–71.

Smith, Michael. 1994. *The Moral Problem*. Oxford: Blackwell.

———. 2003. "Humeanism, psychologism, and the normative story." *Philosophy and Phenomenological Research* 67 (2):460–467.

Turri, John. 2009. "The Ontology of Epistemic Reasons." *Noûs* 43.3:490-512.

Velleman, David. 1989. *Practical Reflection*. Princeton University Press.

Wedgwood, Ralph. 2002. Internalism Explained. *Philosophy and Phenomenological Research* 65 (2):349-369.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford University Press.