

SWIMMING UPSTREAM - PROBLEMS FOR SMITH'S ACCOUNT OF THE NATURE OF REASONS

Jeppje Berggreen Høj

ABSTRACT: In an account of what we might call fundamental practical normativity, it is helpful, I think, to distinguish between the normativity of reasons and that of rationality. But when we do so the question of the relation between these two realms arises: are these two independent kinds of normativity? Can we account for the normativity of rationality in terms of the normativity of reasons? Or is it the other way around; is rationality explanatorily prior to reasons? This paper discusses a positive answer to the last of these questions.

In a number of articles, as well as in his first book, Michael Smith has argued that the nature of reasons can be accounted for in terms of ideal rationality. The image of a fully rational version of an agent, *A*, can help us understand what the claim that *A* has a reason to Φ amounts to. In fact, Smith suggests, claiming that *A* has reason to Φ amounts saying no more or no less than that *A*'s ideally rational self would want *A* (as he is now) to Φ . In what follows, I argue that Smith is unsuccessful in his analysis of normative reasons in terms of full rationality, and that the failure of his analysis leaves his closely related argument for normative judgment internalism unconvincing.

KEYWORDS: Rationality, practical reasons, normative reasons, practical rationality, acting for a reason, normativity, irrationality.

1. Smith on Reasons

According to Michael Smith, what it means for me to have a reason¹ to Φ is the following: a fully rational version of me (my fully rational self) would want me (as I am now) to Φ .² If this is true, it establishes a very close connection between our evaluative beliefs and what we rationally ought to desire – in other words, the model is an analysis indicating an

¹ Smith distinguishes between two kinds of reasons; a) reasons that explain certain action or attitudes and b) reasons that justify action. Here we are concerned with the latter type.

² Smith (1995, 125).

attractively close relation between reason judgements and appropriate motivation. The relation between the two is so close in fact that Smith calls it analytic.³ This, Smith says, is so because if my evaluative belief, that I have reason to Φ , is equivalent to the belief that my fully rational self would desire that I Φ , then not to desire to Φ is by definition less than fully rational. It is to some degree irrational.

This is a very attractive conclusion, he tells us, because there seems to be no other plausible way of tying together our reason judgements with our actions (or at least with our intentions). And surely any account of reasons must be able to explain the close connection between reason judgements and actions, because we take it for granted that such a connection exists.

Smith is surely right that the connection between reason judgements and actions must be accounted for in any plausible account of practical reasons. But it is important to keep in mind exactly what kind of connection this could be. A helpful way to spell this out might be to think of two different readings; a motivational reading and a normative reading. The motivational reading focuses on how the appropriate action comes about when we make reason judgements, that is, what motivates us to act in accordance with our judgements. A general desire to act in accordance with our reason judgements might account for this motivational link, but this seems to Smith,⁴ as to me, an unattractive solution. If we act in accordance with our reason judgements because of such a general desire to do what we have reason to do, we have surely detached ourselves too much from the facts that make it the case that we have reason to act, e.g. that the old lady needs help. When I help the old lady, I do not do so because her need for help instantiates a reason for me to help her and that I want to act in accordance with the instantiated reason – I help her because she needs help. My motivation comes directly from (knowledge of) her need. A general desire to act in accordance with my reason judgements is an unattractive middle link in the motivational connection between my reason judgements and my actions.

A much more attractive way to account for the connection between reason judgements and actions is to ascertain a normative link between the two. Such a link would not try to account for the fact that we are mo-

³ Smith (1995, 109).

⁴ Smith gives an argument along these lines in chapter 3 of Smith (1994).

tivated to act in accordance with our judgements, rather it would explain why we should be. And it is indeed Smith's attempt to demonstrate such a normative link that I shall deal with in this essay.

Before I proceed to dealing with Smith's arguments, I want to bring to the reader's attention how, just as at the motivational level, an unattractive middle link might sneak in at the normative level. In an account of the normative connection between our reason judgments and the appropriately related actions, it is important that we keep in mind precisely what we are trying to connect. When I say that we are investigating the normative connection between our reason judgements and our actions, by the former I do not mean to indicate judgements that explicitly use the concept of a reason. A reason judgement might be: "This lady needs help". And it is this kind of judgement that we are trying to tie together with our actions, and not judgements such as: "I have reason to help the lady" or "The lady's need makes it the case that I have reason to help her". We use these kinds of judgements – let us call them secondary reason judgements – in our speculations about our primary reason judgements (she needs help) and we might perhaps even say that the secondary reason judgements follow from the primary ones. But to repeat, the normative connection we are investigating is between our judgement "she needs help" and our action of helping her (or at least intending to do so) – NOT between our judgement "I have reason to help her" and helping her.

This ends my preliminary remarks, now onto Smith's arguments. I start out with an example of my own.

2. Forcing Rationality into Reasons

I hope that you, the reader, will agree that if swallowing a certain pill will improve my health, then this fact constitutes a reason for me to swallow the pill. And if this is so, a fully rational version of me,⁵ evaluating my situation from a parallel world (Smith calls this an evaluating world) would most likely advise me to swallow the pill. There is in this way, I agree with Smith, clearly a connection between what my fully rational self would advise me to do and what I have reason to do. Of course it is not the fact that my fully rational self would advise me to

⁵ I shall say more below about what it means for someone to be fully rational, according to Smith.

swallow the pill that is my ground for swallowing the pill. My ground for swallowing the pill, what favours my swallowing it, is a feature of the immediate situation, i.e. the fact that swallowing the pill would improve my health.

This is important. Smith is not saying that my grounds for swallowing the pill are that my fully rational self would advise me to do so. This would be absurd as it would leave no role to play for the fact that it would improve my health. My fully rational self would of course advise me to swallow the pill precisely because it would improve my health. That health fact is the ground both for my swallowing the pill and for my fully rational self's advice to do so.

But what is then the relation between the fact that swallowing the pill would improve my health and the fact that my fully rational self would advise me to swallow the pill? Well, Smith describes this relation in different ways. He says:

[...] it is a platitude to say that what it is desirable that we do is what we would desire if we were fully rational (Smith 1994, 150).

And he says that facts about what I have reason to do follow from facts about what it is desirable that I do:

[...] what we have reason to do is equivalent to, [...] or somehow entailed by claims about what it is desirable that we do (Smith 1995, 109).

These quotations show us, I think, that Smith believes that both facts about what we have reason to do and facts about what our fully rational self would want us to do are entailed by facts about what it is desirable that we do. And what it would be desirable that we do has to do with what would be valuable to do, i.e. improve my health.⁶

If this is true, then facts about our reasons are neither prior nor secondary to facts about what we would advise ourselves to do if we were fully rational. Neither one follows from the other. Smith's point is rather this; the fact about my health is what favours my swallowing the pill – and when we talk of reasons to act, it is such favouring relations (between the fact that the pill can improve my health and my action of swallowing it) that we are referring to. But saying that such facts ground

⁶ The introduction to "Internal Reasons" indicates that Smith assumes that such a relation between desirability and value holds true and I am assuming that improvements in health would be a reasonable example of something that has value.

or constitute reasons doesn't tell us much about what it means for a fact to favour an action. Assigning a general name, reasons, to such favouring relations gives no further understanding of what the nature of the relation is. What does give us some understanding of what it means for a fact to favour an act is the image of the fully rational self. Smith is telling us that what it means for me to have reason to swallow the pill, is that my fully rational self would want me to swallow the pill.

In other words, if I were to ask what it means that the fact that swallowing the pill would improve my health favours my swallowing the pill, Smith's account offers the following answer: if this fact pertains then my fully rational self would advise me to swallow the pill. The advice model is offered as a way of explicating the favouring relations that hold between reasons and the actions that those reasons recommend.⁷

But Smith's analysis of the reason relation cannot work. According to Smith, we are to understand the reason relation in terms of what my fully rational self would advise me to do – and I have no other way of understanding the advice that my fully rational self would give me other than assuming that he would advise me in accordance with the reasons that obtain in my circumstances (remember that 'reason' was just the name used for features of the situation that stand in a favouring relation to potential action).⁸ But if my fully rational self would advise me according to the reasons that obtain, then surely a plausible analysis of these reason relations cannot include the advice from my fully rational self. An analysis of reasons in terms of advice grounded in those very reasons is no analysis at all. In order for my reason relations to be analysed in terms of what my fully rational self would advise me to do, his advice to me must be independent of these reason relations themselves. But, as I have argued, this seems impossible. For what would be the grounds for the advice from my fully rational self if not the favouring relations, the nature of which we were trying to understand in the first place?

Another way of saying this is the following: If I were to ask my fully rational self: "why do you want me to help the lady", it seems to me that

⁷ In the introduction to his recently published collection of essays, Smith calls facts about the fully rational self's advisory desires 'truth makers' for claims about normative practical reasons (Smith 2004, 2).

⁸ Remember that this was the only plausible reading of Smith's analysis

he has no answer to give! If he refers to the reasons, then he is already including the concepts we were supposed to analyse!

Now, perhaps someone might object to this criticism along the following lines; even though Smith cannot analyse the concept of a normative reason into ideas about what a fully rational agent would advise, one might still shed some light on the idea of a normative reason by showing the impact normative reasons would have on the desires of a fully rational person. The circularity that seems to threaten Smith's account might not be a vicious circularity but rather a positive circularity, where each of the concepts in the circle helps to explain the nature of the other concepts.

But even this consideration seems mistaken to me. Showing what consequences normative reasons have for the desires of a fully rational person does not seem to me to shed any light at all on the concept of a normative reason. Rather it seems to me to provide us with information about what we mean when we speak of fully rational agents, namely, in this case, a perfectly coherent and informed person with ideal responsiveness to normative features in the world (ideal responsiveness to reasons). What is going on here, then, is the beginning of an account of what it means to be rational in terms of reasons. That is, an account of the relation between reasons and rationality where reasons clearly are explanatorily prior.

But while this is certainly an interesting analysis of the relations between reasons and rationality, the topic for this essay is Smith's attempt to go the opposite way and analyse reasons in terms of ideal rationality.

The overall problem for Smith is this: in his attempts to analyse the notion of a reason in terms of full rationality he is left with two options. The first option is that he can start out from a formal account of rationality, which has to do with relations between one's attitudes. One is rational in these terms if one moves correctly from the attitudes (beliefs, desires, intentions) one actually has, independently of whether one is justified in having these intentional states or whether one's beliefs correspond to the external facts of the world. On this model, rationality has essentially to do with requirements in hypotheticals: "You ought, given that you believe *p*, to believe *q*".

This model of rationality will not get Smith what he wants. If one tries to account for the nature of reasons, in terms of believing, desiring, intending correctly from the attitudes that one happens to have, then

one's reasons become entirely relative to one's intentional states, and these attitudes become reasons for themselves. This is not an attractive result.

The other possibility for Smith is to have a substantive notion of rationality to start out from. Rationality, then, is the formal account from before plus perfect information and, importantly, perfect responsiveness to normative relations in one's environment – so not just descriptive knowledge, but knowledge about what facts favour certain actions.

If Smith's chooses this starting point, he avoids the problems of intentional states justifying themselves – because now he has external justification – but another problem arises. If Smith builds into the notion of rationality, perfect responsiveness to normative relations or features in the environment around him, he has build in the features that he was supposed to account for. For, perfect responsiveness to normative relations is perfect responsiveness to reasons. Smith seems to be saying: what is a favouring relation? Well, it is a relation, which an agent who is fully responsive to those relations (FRS) would respond to. This is uninformative.

And if Smith chooses not to build in responsiveness to normative relations, then the fully rational self has nothing in which to ground his advisory desires. Why should my perfectly rational self desire that I swallow the pill, if he does not know that the health improving facts favour swallowing the pill?

The choice between a formal and a substantive notion of rationality to start out from is, therefore, for Smith, an unpleasant dilemma – each horn of which will impale his attempts to account for reason in terms of full rationality.

So far, I have tried to show why Smith's analysis of the reason relation cannot work. However, even if the advice model cannot account for the nature of the reason relation, this does not guarantee that Smith is wrong about there being a close relation between reason judgements and actions. Even if the nature of my having a reason cannot be accounted for in terms of hypothetical advice from my fully rational self, it might still be the case that my fully rational self would always advise me to act in accordance with my reasons. And if this is the case, then acting against my reason judgements looks less than fully rational. In the next section I will examine whether the relation between the reason judgements and the advice facts really is as close as this. To anticipate, I will conclude that it is not.

3. Smith and Pro Tanto Reasons

There is a great deal of intuitive plausibility to the suggestion that my fully rational self would advise me to act in ways that I have good reason to act. In fact it is often true. But it is not the whole truth, for though my fully rational self would never advise me to do something that I do not have good reason to do, there are other things that I have reasons to do than those which my fully rational self would advise me to do. In other words, the set of things my fully rational self would want me to do is only part of the set of things that I have reason to do. Why is this? Well because sometimes I have reason to do something that I should not do because I have more reason to do something else.

Here is an example of such a situation. The fact that I would enjoy driving a Range Rover is a good reason for me to buy a Range Rover, but this reason is dramatically outweighed by the reasons I have to spend my money on other things, such as food and rent. Smith's account does not immediately seem capable of making sense of the reasons that have been outweighed, because the notion of advice which is so central in Smith's account of normative reasons, necessarily is an 'overall' notion as opposed to a pro tanto concept of reasons. Smith's fully rational self always gives advice regarding what I should do, and cannot identify what things give me reasons to act in particular ways 'as far as this goes'. But we cannot settle for only the overall advice notion that Smith provides. If I were to give myself advice on whether to buy a Range Rover or not, I would not advise myself 'a little' to buy it, and 'a lot' not to buy it. I would only advise myself not to buy it, because this is what I have most reason to do.

But perhaps there is yet another reply available to Smith. My complaint so far has been that if Smith accounts for normative reasons in terms of what he calls the advice model he is unable to account for pro tanto reasons. Advice is an overall notion. But maybe it is worth taking another look at the content of Smith's advice model because actually calling it an advice model is a little misleading (though this is what Smith himself calls it). Here is one of Smith's versions of the advice model:

[the advice model] ... tells us that facts about the desirability of acting in certain ways in the evaluated world are constituted by the desires we would have about the evaluated world in the evaluating world (Smith 1994, 151; my underscore).

The normative reason according to this quotation does not depend on the advice that I would receive from my fully rational self, it depends rather on what my fully rational self would desire that I do.⁹ And while it seems implausible that one can advise oneself a little to Φ and at the same time advise oneself a lot not to Φ , it is quite plausible that one's fully rational self could to some extent desire that one Φ 's and at the same time desire even more that one does not Φ . So now it looks like Smith's "advice" model can fit the notion of a pro tanto reason.

But again there is a problem. If we can make sense of the idea that the fully rational self desires in accordance also with my outweighed pro tanto reasons, we seem to be able to hold onto the idea that the desires of the fully rational agent always track the reason I have. But if we hold this together with the idea that if I do not desire in accordance with my fully rational self, I am irrational, then we get strange implications. Here is an example. I think that I have a pro tanto reason to rob a bank. If I did, I could do some good things with the money for me and my kids (say). If this is so, then if we believe that Smith's claim holds for outweighed pro tanto reasons, my fully rational self would to some extent desire that I rob the bank. And I would be irrational if I did not to some extent want to rob the bank, if I judge myself to have a pro tanto reason to do so. Even when I judge that this reason is dramatically outweighed by other reasons (it would be very bad for me and my kids when I go to jail). This seems counterintuitive to me. I would imagine that a sign that someone is a rational agent would be that she desired in accordance with her overall reason judgements, and not all of them.

This is not, I think, a knock down argument against Smith's claim about the connection between my fully rational self and me as I am now. But I do think that it counts against it. However, even if Smith is wrong about pro tanto reasons, the connection he suggested between my overall reason judgements and the desires of my fully rational self might hold. And that is all he really needs for his claim about overall normative judgements internalism; i.e. the claim that if I judge myself to have most reason to Φ , and yet remain motivationally unaffected by this, then

⁹ This is important for Smith, because if the normative reason depended on what my fully rational self would *advise* me to do, then it would not be irrational for me not to *desire* (and thus intend) to do what I judge myself to have reason to do. It would only be irrational not to *advise* myself to act in accordance with my reasons. And this is not the internalism that Smith is after.

I am irrational. In the last section of this paper, I shall examine whether Smith succeeds in his argument for internalism. Again, as you might by now expect, my conclusion will be that he does not.

4. Consequences for Normative Judgement Internalism

It is important to keep in mind that the kind of internalism that Smith is after is judgement internalism, which says that given that one judges oneself to have reason to Φ , one must be motivated to Φ , on pain of irrationality. Nothing here is said about the relation between one's having a reason to Φ and being motivated accordingly. In speaking of judgement internalism we are, therefore, only concerned with the kind of rationality that I label formal rationality; reacting correctly to the intentional states one has, independently of the status of those intentional states. On this account of rationality, one is practically irrational if one acts contrary to one's own judgements and rational if one acts in accordance with them. With Smith we might call this rationality/irrationality 'by one's own lights'.

The question then is this: given this 'by one's own lights' account of rationality, does the close relation between what one's reason judgements and one's fully rational self's advisory desires make it impossible to remain rational while motivationally untouched by one's reason judgements?

Initially we might think that the answer to this question depends on whether I am making reason judgements of what I above called the primary or the secondary kind. Am I making judgements that include explicit use of the reason concept? We might think that if I am not making judgement about Φ 'ing, explicitly using the reason concept, then whether or not my fully rational self would desire that I Φ , given that I had reason to Φ , is of no consequence for my internal coherence. If I haven't made judgement about reasons, then I cannot possibly be rationally required to act in accordance with any reasons, we might think.

But this would be wrong. We must remember that the advisory desires from our fully rational self, were not based on reason facts. They were based on facts about what it is desirable or valuable that I do, exactly like my primary reason judgements are. So we cannot escape internalism just by not using the concept of a reason in our reason judgements.

Are we then left with no choice but to be internalists? Well not yet, because if we accept a narrow account of rationality – where rationality is defined in terms of coherence – the person who judges himself to have reason to Φ can perfectly coherently refrain from Φ 'ing and still be rational. His is left open only to an accusation of irrationality if he judges both a) that he has reason to Φ and – and this is the crucial bit – b) that if someone has a reason to Φ their fully rational and fully informed self would want him to Φ . Having both these beliefs and still not being motivated to Φ would be straightforwardly incoherent and thus irrational by one's own lights. However, I hope my reader will agree that only very few people (philosophers presumably) would have this rather technical second order belief about his reasons. In other words, if Smith accepts the narrow account of rationality, his internalism will apply only to philosophers. And this is surely not a very attractive version of internalism.

Smith's reply to this criticism would be something like this: if the agent does not believe that if she has reason to Φ then her fully rational self would want her to Φ , this is so only because the agent has a false belief about what it means to have a reason to Φ . For 'I have reason to Φ ' simply means 'my fully rational self would desire that I Φ '. And from the fact that an agent with a false belief about what it means for something to be a reason fails to believe that her fully rational self would want her to Φ when she believes that she has reason to Φ , we should not draw a general conclusion that one can rationally believe oneself to have reason to Φ and not desire to Φ .¹⁰

But this reply from Smith does of course not work. Because as I attempted to show in the first half of the essay, it is a mistake to say that what it means that I have reason to Φ is that my fully rational self would Φ . Even if we think that there is a constant conjunction between the two, one can surely rationally believe one without the other.

End Note

In this paper I have dealt with two closely related suggestions from Smith: a) that normative reasons can be accounted for in terms of what my fully rational self would want me to do. And b) that it follows from

¹⁰ Smith (2001, 263).

the nature of normative reasons that any agent who does not act according to her reason judgements is irrational by her own lights. I have argued that the first suggestion is unattractive as Smith is left with an impossible choice between two accounts of what it means to be rational and that the second falls as a consequence of this; given that reasons cannot be analysed in terms of formal rationality, the normative ties between reason judgements and motivation are too loose to sustain judgement internalism.

The University of Reading
Whiteknights, PO Box 217
READING, Berkshire
RG6 6AH
United Kingdom
jeppé.hoej1@gmail.com

REFERENCES

- SMITH, M. (1994): *The Moral Problem*. Oxford: Blackwell.
- SMITH, M. (2004): *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-Ethics*. New York: Cambridge University Press.
- SMITH, M. (2001): The Incoherence Argument: Reply to Schafer-Landau. *Analysis* 61, 3, 254 – 266.
- SMITH, M. (1995): Internal Reasons. *Philosophy and Phenomenological Research* 55, 1, 109 – 131. Reprinted in Wallace, R. J. (ed.) (1999): *Reason, Emotion and Will*. Aldershot: Ashgate Publishing, 161 – 183.