

## IS THERE A NEXUS BETWEEN REASONS AND RATIONALITY?

Michael Smith

When we say that a subject has attitudes that she is rationally required to have, does that entail that she has those attitudes for reasons? In other words, is there a deep nexus between being rational and responding to reasons? Many have argued that there is. For example, Derek Parfit tells us that 'to be rational is to respond to reasons' (Parfit 1997, p.99). But I am not so sure. I begin by considering this question in the domain of theoretical rationality. The question in this domain is whether, when a subject has the beliefs that she is required to have by the norms of theoretical rationality, she is responding to reasons that there are for having those beliefs. Armed with a moderately clear answer to this question in the theoretical domain, I consider their relationship in the practical domain. When a subject has the desires that she is required to have by the norms of practical rationality, is she responding to reasons that there are for having those desires? Part of the interest of these questions lies in improving our understanding of reasons for action. I will say a little about this towards the end.

### 1. Reasons and rationality in the theoretical domain

Let's begin with a very simple case of theoretically rational belief formation. Suppose that the following is true:

TR: Reason requires that if I believe that p and I believe that if p then q, then I believe that q

Furthermore, suppose that I believe that p, and that I believe that if p then q, and that, abiding by TR, I rationally go on to form the belief that q — let's just assume whatever else needs to be true for this to be the case. Should we conclude that, in such circumstances, there must be reasons for the formation of the belief that q? In

other words, does the mere fact that I am theoretically rational in the formation of the belief that q entail that there are reasons to which I am responding?

It is important, in answering this question, that we do not move unwittingly between different senses of the term 'reason'. There is a widely accepted distinction in the philosophical literature between two senses in which we talk of reasons for action (Woods 1972, Smith 1987). On the one hand, some of our talk of reasons is talk about psychological states that are capable of rationally explaining an action. On the other, some of our talk of reasons is talk, not of psychological states that explain, but about considerations that justify. In these terms, the concern is that there is a similar distinction to be made in our talk of reasons for belief. Specifically, what we have asked appears to be ambiguous between a question about reasons for belief in the sense of psychological states that explain our beliefs, and reasons for belief in the sense of considerations that justify our beliefs. In answering the question we must therefore make explicit the sense of the term 'reason' that we have in mind.

Let's begin by considering reasons in the sense of psychological states that rationally explain. It is uncontroversial that, when I am theoretically rational in the formation of the belief that q in circumstances like those described, there are psychological states that explain my believing that q: I believe that q because I believe that p and I believe that if p then q, and it is in virtue of there being such a psychological explanation that I count as being theoretically rational. Since the nature of this psychological explanation will be important in what follows, let's say that in such circumstances the beliefs that p and that if p then q evidentially explain the belief that q. If the issue is whether there is a nexus between reasons in this sense and being theoretically rational then, in the circumstances described, the answer is that there is.

But what about reasons in the other sense? If, in circumstances like those described, I were asked what my reasons are for forming the belief that q, in the sense of the

considerations that justify my belief, then I would unhesitatingly insist there are such reasons, namely, that  $p$  and that if  $p$  then  $q$ . But this doesn't yet entail that there are in fact reasons for my forming the belief that  $q$ , still less that the considerations I cite are reasons for doing so. Whether or not this is so turns on the relationship between the reasons that there are for forming the belief that  $q$  and the reasons that I would cite as my reasons for forming that belief. Let's therefore consider that relationship.

When I cite the facts that  $p$  and that if  $p$  then  $q$  as my reasons for believing that  $q$ , it seems to me that I do so in full recognition of the possibility of there being a gap between the reasons that there are for believing that  $q$  and the reasons that I would cite. For one thing, I could readily admit that I have other reasons for believing that  $q$ , reasons that I don't presently know about. For another, I could readily admit that what I take to be reasons for believing that  $q$  are no such thing: that though I think that there are considerations that justify my believing that  $q$ , there aren't really.

As regards the first, it seems to me that I could happily admit the possibility of evidence available to me, evidence to which I am not currently responsive, that supports  $r$ , where if  $r$  then  $q$ . Moreover, if there were such evidence, then it seems to me that my reasons for believing that  $q$  would be more expansive than I currently think. Though I think that the only reasons I have for believing that  $q$  are that  $p$ , and that if  $p$  then  $q$ , I can thus accept the possibility that I have other reasons for forming that belief as well, namely, that  $r$ , and that if  $r$  then  $q$ .

This suggests that when I say that certain considerations are my reasons for believing that  $q$ , the 'my' just signals the availability of the reasons in question to me. I might just as well have said that the reasons for believing that  $q$ , those that happen to be available to me, are that  $p$  and that if  $p$  then  $q$ . Put this way, it is clear that what I say in the imagined situation might at best be a partial truth.

As regards the second point, it also seems that when I am asked what my reasons are for forming the belief that  $q$ , I could happily admit that what I take to be reasons for believing that  $q$  are no such thing. If, for example, it isn't the case that  $p$ , or if it isn't the case that if  $p$  then  $q$ , then I would say that though it seems to me that there are reasons for believing that  $q$ , there aren't really such reasons— or, anyway, that no such reasons are known to me.

This might be disputed. For how could the mere fact that my beliefs that  $p$ , or that if  $p$  then  $q$ , are false suffice to show that there is no reason for me to believe that  $q$ ? But the very temptation to ask this question seems to me to depend on assuming what is at issue, namely, that when I theoretically rational in believing what I believe, I believe what I believe for reasons. To see why this assumption is false, we need to reflect a little on the relationship between the reasons that there are for forming beliefs and the beliefs that we form on the basis of those reasons.

If my reasons for believing that  $q$  are that  $p$  and that if  $p$  then  $q$ , and if I believe that  $q$  for those reasons, then it would seem to follow that these reasons don't just justify my belief, but that they also figure in an explanation of my belief: I believe that  $p$  because  $p$  and if  $p$  then  $q$  (compare Williams 1980, p.102). This is, if you like, what the 'for' signals in the claim that I believe that  $q$  for those reasons. But, since all explanation is factive, it follows that, if these are indeed reasons for believing that  $q$ , then they must pick out ways that things really are. In other words, my reasons for believing that  $q$  are that  $p$  and that if  $p$  then  $q$ , then it must be the case that  $p$  and that if  $p$  then  $q$  (Parfit 1997, p.99; though contrast Dancy 2000, pp.131-137).

This suggests that there is an important connection between the two sorts of reasons in the theoretical domain: that is, the reasons in the sense of the considerations that justify and the reasons in the sense of the psychological states that explain. Since, when there are considerations that justify a subject's forming the beliefs he forms, and

the subject forms his beliefs on the basis of those considerations, the considerations themselves must figure in an explanation of his belief, and since, in such circumstances, there must also be psychological states that explain his belief, it follows that the two explanations of the subject's beliefs — that in terms of the considerations and that in terms of the psychological states — must somehow dovetail. And indeed they do. For what happens in such circumstances is this: the fact that  $p$  explains the subject's believing that  $p$ , and the fact that if  $p$  then  $q$  explains the subject's believing that if  $p$  then  $q$ , and this pair of beliefs explains the subject's believing that  $q$ . We will return to the connection between reasons in the sense of considerations that justify and reasons in the sense of psychological states that rationally explain, presently.

The two points just made suggest a preliminary story about the nature of reasons for belief, in the sense of considerations that justify. According to this story, the reasons that there are for believing that  $q$  — the considerations that justify believing that  $q$  — are ways things are which make it rational for someone who believes that things are that way to believe that  $q$ . A subject's reasons for believing that  $q$  — or, equivalently, the reasons a subject has for believing that  $q$  — are a sub-set of these reasons. They are those reasons for believing that  $q$  that are available to that subject. The reasons for which a subject believes that  $q$  are a subset of this subset. They are those reasons for believing that  $q$  that are not just available to that subject, but which the subject is aware of and which figure in an explanation (of the right kind) of that subject's believing that  $q$ .

If some such story as this is along the right lines then it follows immediately that, if there is a nexus between reasons for belief, in the sense of the considerations that justify belief, and rationality at all, then that nexus is nowhere near as deep as we might at first have thought it was. There may be reasons for believing that  $q$  which no one knows about, and perhaps even that no one could know about. So though there

are reasons for believing that q, it may be the case that no one is in a position rationally to form the belief that q. Conversely, even if various people are in a position rationally to form the belief that q, in so doing they might not be responding to reasons that there are to form that belief; they might not be believing that q for reasons. For though the rationality of their belief that q is guaranteed by the fact that their belief conforms to certain norms of theoretical rationality, the evidence available to them might be wildly misleading. There may be no reasons to believe that q, even though they are quite rational in forming that belief, and plainly do so for reasons in the other sense, that is, in the sense that there are psychological states that evidentially explain their belief.

With these observations in place we are in a position to answer the question with which we began. When a subject has the beliefs that she is required to have by the norms of theoretical rationality, it need not be the case that she is responding to reasons for having those beliefs in the sense of responding to considerations that justify. Such reasons for belief are, to repeat, ways things are which make it rational for someone who believes that things are that way to believe the thing in question. But since subjects can have evidentially well supported but wildly false beliefs, they are capable of rationally forming beliefs for no reason in this sense. What does seem to be true, of course, is that when a subject has the beliefs that she is required to have by the norms of theoretical rationality then she responds to what seem to her to be reasons to have those beliefs, in the sense of considerations that justify, and hence her beliefs are rationally explained — or better, evidentially explained — by reasons, in the sense of psychological states. But, to repeat, this falls short of the requirement that there be any reasons for her beliefs, in the sense of there being ways things are that in fact justify her beliefs.

## 2. Is there a nexus between reasons and rationality in the practical domain?

Let's now consider the relationship between reasons and rationality in the practical domain. We will focus on what many take to be a paradigm case of practical rationality, namely conformity to the means-ends norm (Smith forthcoming).

Suppose that

PR<sup>ME</sup>: Reason requires that if I desire to  $\phi$  and I believe that I can  $\phi$  by  $\psi$ -ing then I desire to  $\psi$

Furthermore, suppose that I desire to  $\phi$ , and that I believe that I can  $\phi$  by  $\psi$ -ing, and that, abiding by PR<sup>ME</sup>, I rationally go on to form the desire to  $\psi$  — again, let's just assume whatever else needs to be the case for this to be true. Should we conclude that, in such circumstances, since I conform to the norms of practical rationality, there must be reasons for the formation of the desire to  $\psi$ ?

In answering this question it is once again important that we do not move in a slapdash fashion between two different senses of the term 'reason'. As before, the question we have asked is ambiguous between a question in the practical domain about reasons for desire in the sense of psychological states that explain our desires, and a question about reasons for desire in the sense of considerations that justify our desires.

Let's begin by focussing on the psychological states that explain desires. Once again, it seems that when a subject manifests the fact that she is practically rational by forming the desire to  $\psi$  in the circumstances described, it does indeed follow that there must be reasons why she desires to  $\psi$  in the sense of psychological states that figure in a rational explanation of her desire: she desires to  $\psi$  because she desires to  $\phi$  and believes that she can  $\phi$  by  $\psi$ -ing, and her being practically rational is crucially dependent on there being an explanation of her desire of just this kind. If the issue is whether there is a nexus between reasons in this sense and rationality in the circumstances described, then the answer is that there is. The situation is exactly the

same as in the theoretical domain. It is, however, worth pausing to consider a difference in the nature of the rational explanations in these two domains.

Because the psychological states involved are quite different — beliefs, in the theoretical domain, pairs of desires for ends and beliefs about means to those ends, in the practical domain — it follows that the kinds of psychological explanation that we find in the two domains is quite different as well. In the theoretical domain we saw that the psychological states that rationally explain beliefs explain them in the sense of evidentially explaining them. But the psychological states that rationally explain desires, at least in the circumstances described, do not evidentially explain those desires (whatever that might mean — we will return to this presently). Instead they figure in what appears to be a straightforward teleological explanation of such desires.

To repeat, as we have just seen, there is indeed a nexus in the practical domain between reasons, in the sense of psychological states that rationally explain, and rationality. The question to ask next is whether there is a nexus between reasons and rationality in the other sense of 'reason'. When an agent is practically rational in circumstances like those described, must she be responding to reasons in the sense of considerations that justify?

What would I say if I was asked what my reasons are for forming the desire to  $\psi$  in circumstances like those described? In other words, what considerations would I cite in giving a justification of my desire? Since one of my beliefs figures in the explanation of my desire to  $\psi$  — since a belief of mine is one of the reasons, in the other sense of 'reason', for my desire — it follows that there is at least one consideration that I could cite, namely, the fact that I can  $\phi$  by  $\psi$ -ing. But we should not assume that I take this consideration to be any part of a set of considerations that justifies my desire to  $\psi$ . For the truth is that I would be stuck if I was asked what the justification of my desire is. The question assumes that I believe there to be

considerations that justify my formation of the desire to  $\psi$ , whereas, at least as described, it seems that I need have no such belief. Worse still, on certain assumptions, there could be no such beliefs.

What would it be for there to be a consideration that justifies my desire to  $\psi$ ? There would have to be a consideration that stands to my desire to  $\psi$  in much the same relation as the considerations that justify my beliefs stand to the beliefs that they justify. As we saw above, such reasons for believing that  $q$  are the ways things are which make it rational for someone who believes that things are that way to believe that  $q$ . So, by analogy, reasons in this sense for desiring to  $\psi$  — the considerations that justify desiring to  $\psi$  — would have to be the ways things are which make it rational for someone who believes that things are that way to desire to  $\psi$ .

But if this is right then the problematic nature of such reasons is readily apparent. If the norms of practical rationality are all like  $PR^{ME}$ , the norm requiring us to have desires for means when we have desires for ends and beliefs about the means to those ends, in the crucial respect of being requirements for us to have certain desires, given that we have certain other desires, then there is no way the world is such that someone's believing the world to be that way makes it rational for them to desire anything. Beliefs, all by themselves, are never enough to make it rational for people to have particular desires.

The conclusion is thus the extreme one that if all of the norms of practical rationality are like the means-ends norm in being requirements for us to have certain desires, given that we have certain other desires, then, when a subject has the desires that she is required to have by the norms of practical rationality, she is not responding to reasons for having those desires. There are no such reasons, and it is no part of our conception of what it is to be practically rational that agents take there to be such

reasons either. Reasons in the sense of considerations that justify desires have nothing whatsoever to do with being practically rational.

### 3. Two ways of resisting the extreme conclusion in the domain of practical rationality

In order to resist the extreme conclusion that reasons for desires, in the sense of considerations that justify having desires, have nothing to do with being practically rational, we would have to reject the assumption on which it is premised. The assumption is that the norms of practical rationality are all like the means-ends norm in being requirements for us to have certain desires, given that we have certain other desires. But on what grounds could we reject this assumption? There are two main strategies: the Besire Strategy and the Rationalist Strategy.

The Besire Strategy turns on the possibility of there being beliefs whose possession entails the possession of certain desires. Assume that  $PR^{ME}$  is true, and assume in addition that some claim of the following form is true:

BD: If an agent believes that  $p$  then she desires to  $\phi$ .

With these assumptions in place we can derive the following principle:

$PR^{BD}$ : Reason requires that if I believe that  $p$  and I believe that I can  $\phi$  by  $\psi$ -ing then I desire to  $\psi$

$PR^{BD}$  is a principle telling us which desires we are required to have given that we have certain beliefs. To be sure, the full story behind the truth of  $PR^{BD}$  goes via  $PR^{ME}$ , which is a principle telling us which desires to have given that we have certain other desires, and B. But though this is an important fact about  $PR^{BD}$ , one which will be important to remember presently, it is plainly consistent with this that  $PR^{BD}$  is not itself – as it quite evidently is not! — a principle that tells us which desires to have,

given that we have certain other desires.  $PR^{BD}$  tells us which desires to have on condition that we have certain beliefs.

But now suppose that  $PR^{BD}$  is true. It then follows that there could indeed be considerations which justify the formation of certain desires. For suppose that  $p$  and that I can  $\phi$  by  $\psi$ -ing. There is then a way the world is, namely, the way characterised by ' $p$ ' and 'I can  $\phi$  by  $\psi$ -ing', which makes it rational for me to have a certain desire, namely the desire to  $\psi$ , on condition that I believe that the world is indeed this way.

Nor is it hard to think of values for ' $p$ ' and ' $\phi$ ' which, at least according to some theorists, makes a principle like  $PR^{BD}$  come out true. John McDowell has argued that when we believe that some person,  $A$ , is shy and sensitive — believe in the sense of fully understanding what it is for  $A$  to be shy and sensitive — then our having this belief entails our desiring to treat  $A$  in certain ways: to (say) protect  $A$  from those who would exploit her vulnerability (McDowell 1978). Let's call this

$BD^{McD}$ : If an agent believes that another person,  $A$ , is shy and sensitive then his having this belief entails his desiring to protect  $A$  from those who would exploit her vulnerability.

Suppose McDowell is right about  $BD^{McD}$ . Then, given  $PR^{ME}$ , we can derive the following norm of reason governing the formation of desires on the basis of beliefs:

$PR^{BD/McD}$ : Reason requires that if I believe that  $A$  is shy and sensitive and I believe that I can protect her from someone who would exploit her vulnerability by  $\psi$ -ing then I desire to  $\psi$

So if what McDowell's says about the nature of the belief that someone is shy and sensitive is true, then there are considerations that justify the formation of certain desires to perform particular acts: that certain people are shy and sensitive and that a certain agent's acting in a particular way would protect them from those who would

exploit their vulnerability are considerations that justify that agent's desiring to act in that particular way. These are reasons for the agent to desire to act in that way because they are considerations which make it rational for him to desire to act in that way on condition that he believes that those considerations obtain.

It is important to note that the plausibility of the Besire Strategy as such isn't tied to the plausibility of McDowell's own view. The crucial question is not whether the belief that someone is shy and sensitive has the character McDowell claims it has, but rather whether there are any beliefs whose possession entails the possession of certain desires. If there are then, along the lines just sketched, we will be able to show that there could be reasons for the agent to have desires.

An alternative way to reject the assumption that the norms of practical rationality are all requirements for us to have certain desires, given that we have certain other desires, is to pursue the Rationalist Strategy. The Rationalist Strategy grants that there are no beliefs that entail the possession of particular desires. Instead it makes a normative claim. It tells us that there are additional basic norms of practical rationality, over and above  $PR^{ME}$ , norms of the following form that require us to have certain desires, given that we have certain beliefs.

$PR^R$ : Reason requires that if I believe that p, then I desire to  $\phi$ .

Unlike  $PR^{BD}$ ,  $PR^R$  allows that an agent can have the belief that p without having the desire to  $\phi$ . It simply tells us that that combination of belief and desire violates a norm of reason.

If some version of  $PR^R$  is true then it would once again follow that there could be considerations which justify the formation of certain desires. For suppose that p. There is then a way the world is, namely, the way characterised by 'p', which makes it

rational for me to have a certain desire, namely the desire to  $\phi$  on condition that I believe that the world is indeed this way.

Nor is it hard to think of a value for 'p' which, at least according to some theorists, makes a principle like  $PR^R$  come out true. According to one standard reading of Thomas Nagel's The Possibility of Altruism (1970), for example, Nagel argues that if someone believes that another creature is a person, equally real as himself, and he believes that that other person is in pain, and he believes that he can relieve that person's pain by acting in a particular way, then he violates a norm of practical reason if he doesn't desire to act in that particular way (Wallace 1990; for an alternative reading of Nagel according to which he is pursuing the Besire Strategy, see Smith 1994, Ch.4). The following is thus, by Nagel's lights, a self-standing norm of practical rationality alongside the means-ends norm:

$PR^{R/N}$ : Reason requires that if I believe that another creature, A, is a person, equally real as myself, and I believe that A is in pain, and I believe that I can relieve A's pain by  $\phi$ -ing, then I desire to  $\phi$ .

So, if what Nagel says is true, then there are considerations that justify the formation of certain desires to perform particular acts: that another creature is a person, equally real as an agent himself, and that that person is in pain, and that the agent himself can relieve that person's pain by acting in a particular way, are considerations that justify the agent's desiring to act in that particular way. These are reasons for the agent to desire to act in that way because they are considerations which make it rational for him to desire to act in that way on condition that he believes that those considerations obtain.

As with the Besire Strategy, it is important to note that the plausibility of the Rationalist Strategy as such isn't tied to the plausibility of Nagel's view. The crucial question is not whether the beliefs Nagel cites have the characteristics he says they

do, but rather whether there are any beliefs whose possession makes it rational for an agent to have certain desires. If there are then, along the lines just sketched, we will be able to show that this entails the possibility of reasons for the agent to have desires.

Though the two strategies both underwrite the possibility of there being reasons for desires, in the sense of considerations that justify having desires, as we have already noted in passing, there is a significant difference in the way in which the two strategies underwrite this possibility. This difference underscores a difference in the two strategies' views about the nature of the psychological states that explain desires, and hence a difference in what the two strategies tell us about the nature of reasons for desires in the other sense of 'reason', the sense in which reasons are psychological states that explain desires.

To repeat, the Besire Strategy tells us that the desires that are justified are produced in the normal means-end way:  $PR^{B/McD}$  is derived from  $BD^{McD}$  and  $PR^{ME}$ . The agent has a belief that entails that he has a desire for an end, and this desire for the end combines with a belief about a means to that end to produce a desire for the means in the normal means-end way. The desire that is justified — the desire for the means — is thus susceptible to a regular teleological explanation. The reason for the desire, in the sense of the psychological state that explains that desire, is thus a psychological state that is suited to play the required role in this teleological explanation. It is a desire, albeit a desire whose possession happens to be entailed by the possession of a belief.

However the Rationalist Strategy, by contrast, tells us that desires that are justified are produced in a completely different way:  $PR^{R/N}$  is supposed to be a basic norm of practical rationality alongside  $PR^{ME}$ . The agent has a belief which produces the relevant desire, but this belief does not produce that desire by combining with a desire for an end. The desire that is justified is thus not susceptible to a regular teleological explanation, and the reasons for the desire, in the sense of the psychological states that

explain the desire, that are posited by the Rationalist Strategy are thus not psychological states that are well suited to providing a teleological explanation. But in that case, what sort of explanation is it?

It is irresistible, I think, to conclude that the reasons for desires, in the sense of the psychological states that explain desires, posited by the Rationalist Strategy are supposed to explain desires in a way that is strongly analogous to the way in which the reasons, in the psychological state sense, for belief explain beliefs. In other words, notwithstanding the fact that it is an explanation in the practical domain, the style of explanation is strongly analogous to evidential explanation. Of course, the explanation is not evidential explanation pure and simple, for the beliefs in question do not support the truth of the desire: desires aren't the sort of state that can be true or false. But the explanation is strongly analogous to evidential explanation in so far as the beliefs in question suffice to make the desire in question the one that it is sensible to have independently of what other desires are had. In this extended sense the beliefs bear evidentially on the desires: the beliefs mandate possession of the desires all by themselves; to desire otherwise is to fly in the face of the facts.

#### 4. Is either of the ways of resisting the extreme conclusion plausible?

Is either of these strategies plausible, and, if so, which is more plausible? Though this is not the place to argue the point at great length, my own view is that the extreme conclusion is unstable. Even if we begin just by assuming that norms of practical reason tell us which desires to have, given that we have certain desires, we are quickly led from this to the conclusion that there are norms telling us which desires to have, given that we have certain beliefs. Moreover, we are led to this conclusion in the manner suggested by the Rationalist Strategy, not the Besire Strategy, for the Besire Strategy is hopeless.

The problem with the Besire Strategy, in a nutshell, is that it requires a far too demanding account of what it takes to understand a proposition. In many ways, these flaws are analogous to flaws in the view of belief according to which belief is closed under a priori consequence. Since I take it that few people hold that view about belief, let me bring out its flaws first, and then spell out the analogy. It will then be clear not just why we should reject the Besire Strategy, but why we should focus our attention on the Rationalist Strategy.

Suppose  $p$  a priori entails  $q$  and consider the principle:

BB: If a subject believes that  $p$  then she believes that  $q$ .

This principle is immensely implausible, on the face of it, for the fairly flat-footed reason that it is one thing to understand and believe that  $p$  and quite another to see that  $q$  is an a priori consequence of  $p$ . An example will help bring out this flat-footed point.

Mathematicians debate long and hard about whether various mathematical conjectures are true or false. But, if true, these conjectures are, let's suppose, a priori consequences of various other more basic mathematical propositions that the mathematicians claim to believe. Now suppose that a particular mathematical conjecture is true. Should we conclude that the mathematicians who claim to believe these more basic mathematical propositions are either speaking truly, in which case they already believe the conjecture, notwithstanding the fact that they claim that they don't know whether the conjecture is true or false, or that they are speaking falsely because, since they are right that they don't know whether the conjecture is true or false, it follows that they mustn't really understand the more basic mathematical propositions from which the conjecture follows a priori? If, like me, you find it extremely hard to believe either of these things then you have no choice but to agree that BB is implausible.

But what is the alternative to BB? The alternative is to suppose that, when p is an a priori consequence of q, the belief that p and the belief that q stand in the following normative relation.

$TR^{APC}$ : Reason requires that if a subject believes that p then she believes that q.

$TR^{APC}$  is a principle of theoretical rationality which tells us that though it is possible for someone to believe that p without believing that q when p is an a priori consequence of q, someone who has this pair of beliefs violates a norm of reason. Mathematicians who believe the more basic mathematical propositions from which a true conjecture follows a priori violate a norm of reason in failing to believe the conjecture. This is a very plausible claim.  $TR^{APC}$  is thus far more plausible than BB.

The analogy between the view that belief is closed under a priori consequence and the view of the relationship between belief and desire propounded by those who pursue the Besire Strategy is, I hope, already apparent, but in case it isn't, let me spell it out. The following two principles are extremely similar in crucial respects:

BB: If a subject believes that p then she believes that q.

BD: If an agent believes that p then she desires to  $\phi$ .

For both BB and BD require us to have extremely high standards of what it takes to understand a proposition. And the following two principles are also extremely similar in crucial respects:

$TR^{APC}$ : Reason requires that if a subject believes that p then she believes that q.

$PR^R$ . Reason requires that if an agent believes that p then she desires to  $\phi$

For just as reflection on the inadequacy of BB forces us to focus on the more plausible  $TR^{APC}$ , so reflection on the inadequacy of BD forces us to focus on the more plausible  $PR^R$ .

But how plausible is  $PR^R$ ? Are there any values of 'p' and 'ϕ' that make  $PR^R$  come out true? My own view is that there are. Suppose, pro tem, that

$PR^{ME}$ : Reason requires that if I desire to ϕ and I believe that I can ϕ by ψ-ing then I desire to ψ

is the only norm of practical reason. Now suppose that an agent believes that, if she had a desire set that conforms perfectly to all of the norms of reason, then she would desire that she ψs, but that she does not desire to ψ. It seems to me that we can now provide an argument for the claim that the following is a norm of reason

$PR^{R*}$ : Reason requires that if an agent believes that she would want that she ψs if she had a desire set that conforms to all of the norms of reason, then she desires that she ψs.

and, hence, that  $PR^{ME}$  is not the only norm of reason. Moreover,  $PR^{R*}$  is not just an additional norm of reason, it is a norm of reason of the same form as  $PR^R$ . In other words, it is a norm which requires agents to have certain desires, given that they have certain beliefs. So, starting with just the assumption that  $PR^{ME}$  is a norm of reason, we see that, given agents can form beliefs about what they would want if their desires conformed to norms of reason, there are indeed norms of reason of the same form as  $PR^R$ . Let me say a little in support of these claims.

For  $PR^{R*}$  to be plausible, it is important that we focus on the right logical form of the belief in question. The idea is not that the agent believes that, in the nearest possible world in which her desire set conforms perfectly to the norms of reason, she desires that she ψs in that world. Reason certainly doesn't ban an agent's failing to desire to do in this world, in which she does not conform to all of the norms of reason, something that she believes that she would desire herself to do in that world, in which she does. For the believed difference in her circumstances — the fact that she

believes that she conforms to the norms of reason in that world, but not in this world — might well make all the difference to what it would be sensible for her to want in that world as opposed to this. Rather, the idea is that the agent in question believes that, in the nearest possible world in which her desire set conforms perfectly to the norms of reason, she desires that she  $\psi$ s in this world.

Consider an example. Suppose I believe that I desire the creature comforts, and that I can have the creature comforts by earning money, but I also believe that, because I am means-end irrational — remember,  $PR^{ME}$  is the only norm of practical reason that we are admitting at this stage — I don't desire to earn money. Moreover suppose that my belief that I don't desire to earn money is true. In that case, when I consider what I would want myself to do in this world in the nearest possible world in which my desire set conforms perfectly to the norms of reason, I conclude that what I would want myself to do in this world is earn money. I come to this conclusion because, in that world — the world in which my desires conform to all of the norms of practical reason — my desires conform to  $PR^{ME}$ . So the upshot is that I believe that I would desire myself to earn money in this world if I had a desire set that conformed perfectly to the norms of reason, but that I don't desire to earn money in this world.

To repeat, it seems to me that reason is against this pairing of belief and desire. The explanation of this ban is relatively simple. I am, after all, a creature who is not just capable of having beliefs and desires that are subject to norms of reason, but a creature who is also capable of forming views about which beliefs and desires I would have if my beliefs and desires conformed to norms of reason, and a creature who is capable, as well, of managing my beliefs and desires in the light of the views I form. In these terms, the problem with my lacking a desire to earn money in the circumstances described is that it shows the extent to which I fail in that management role. It shows that I fail in that management role in an exactly analogous way to the way in which I would fail in that management role if I failed to believe that  $q$  when I

believe that I would believe that  $q$  if I had a belief set that conforms to all of the norms of theoretical reason (for an alternative view see Sayre-McCord 1997).

If this is right, then it follows immediately that, even beginning from a very minimal assumption about the nature of the norms of practical reason, an assumption congenial to the extreme view, the assumption that  $PR^{ME}$  is a norm of practical reason that governs desires, once we fully internalise the consequences of the fact that we can form beliefs about what we would want if our desires conformed to this norm, and the fact that we can manage our desires in the light of these beliefs, we see that there is a more substantive norm governing our desires,  $PR^{R*}$ , which is a norm of the form  $PR^R$ . The upshot is thus that we must reject the extreme view described earlier. We have no alternative but to pursue the Rationalist Strategy.

##### 5. What have we learned about reasons for action?

I said at the outset that part of the aim of this paper is to improve our understanding of reasons for action. In this final section I will say a little about what we have learned from the previous discussion.

As I have already said, it is widely accepted that we can distinguish two senses in which we talk of reasons for action. Some of our talk of reasons for action is talk about the psychological states that are capable of rationally explaining actions, and some is talk, not of psychological states that explain, but of considerations that justify actions. Let's call reasons of the first sort 'motivating' reasons, and reasons of the second sort 'normative' reasons (Smith 1994, Chapter 4).

As is no doubt obvious, the argument just given for  $PR^{R*}$  is, in effect, an argument in favour of the conception of normative reasons proposed by Bernard Williams in his famous paper 'Internal and External Reasons' (Williams 1980; Pettit and Smith forthcoming). One consequence of our discussion, then, is that we must suppose that

there are at least some normative reasons. Though the extreme view that there are no considerations that justify desires looks like it might well have implied that there are no normative reasons for action either, that extreme view is unstable. There is at least one consideration that justifies desire, namely, the fact that the desirer herself would want herself to act in the way desired if she had a desire set that conforms to all of the norms of practical reason. Whether there are further considerations that justify desires is an open question, one which I hope to pursue elsewhere.

As regards motivating reasons, the standard theory is the one we have inherited from Hume, a theory which has more recently been popularised by Donald Davidson (Hume 1740, Davidson 1963). According to this theory, an agent's motivating reasons rationally explain his action in a characteristic teleological manner: motivating reasons embody the goals that the agent has in acting and they embody his conception of what he is doing as something that will achieve his goals. Hume tells us that motivating reasons are psychological pairs comprising an agent's desires for ends (these embody his goals) and means-end beliefs (these embody his conception of what he is doing as something that will achieve his goals), where desire and belief, in turn, are distinct existences: no matter what combination of belief and desire an agent has, we can always imagine an agent who has those beliefs and yet who has different desires, and vice versa. Moreover, the fact that we can always find motivating reasons — that is, desire and belief pairs — that teleologically explain an action is taken by the standard theory to be constitutive of action: what makes an event an action is the fact that it is something that is done for a reason, in the sense that there is a motivating reason, a desire and means-end belief pair, that teleologically explains it.

It perhaps goes without saying that the extreme conclusion mooted above sits happily alongside the Humean theory of motivating reasons. The extreme conclusion, you will recall, was that though there may be reasons for desires in the sense of psychological states that rationally explain them — desires can, after all, be

teleologically explained by other desires — there can be no reasons for desires in the sense of considerations that justify them. This sits happily alongside the Humean theory of motivating reasons because it simply adds a detail to that theory. It tells us that though there may be considerations that justify the means-end beliefs that partially constitute motivating reasons, there are no considerations that justify the desires for ends that partially constitute motivating reasons. But what if we were to reject that extreme conclusion? Suppose we could successfully pursue either the Besire Strategy or the Rationalist Strategy. Would this force us to adopt a different view about the nature of motivating reasons?

The Besire Strategy challenges a core element in the standard Humean theory of motivating reasons, for it takes issue with the idea that belief and desire are distinct existences. Moreover, because it takes issue with this core idea it must reject the account of what makes an action an action. Those actions that are teleologically explained by desires whose possession is entailed by beliefs the agent has are plainly not actions in virtue of being susceptible to a teleological explanation by desires that are distinct from beliefs. So if we can successfully pursue the Besire Strategy, then we must reject the standard Humean theory of motivating reasons. But since, as I have already explained, the Besire Strategy is hopeless, this challenge doesn't seem to me to be of much concern.

The Rationalist Strategy, by contrast, provides no real challenge the standard Humean theory of motivating reasons. For it takes the idea that belief and desire are distinct existences for granted and simply adds a detail. According to the Rationalist Strategy, the desires that partially constitute motivating reasons can themselves be rationally explained by beliefs, rationally explained in a sense strongly analogous to evidential explanation. Though Hume himself didn't believe that desires were susceptible to such rational explanation, it is plainly consistent with the Humean theory of motivating reasons, as outlined above, that Hume was wrong about this and desires

are susceptible to such rational explanation. Moreover it is also plainly consistent with the possibility of giving such a rational explanation of desire that the fact that this is so is in no way essential to an action's being an action. What makes an action an action is the fact that it can be teleologically explained by a desire for an end and a belief about means, something we can establish to be so under a conspiracy of silence about the rational etiology of that desire and belief. Whether or not we can successfully pursue the Rationalist Strategy is thus orthogonal to the Humean theory of motivating reasons.

### Conclusion

We are now in a position to answer the quite general question that we asked at the beginning of this paper. When we say that a subject has attitudes that she is rationally required to have, does that entail that she has those attitudes for reasons? In other words, is there a deep nexus between being rational and responding to reasons?

The answer to this question is a quite decisive 'No'. A subject's having the beliefs that she is required to have by the norms of theoretical rationality entails, at most, that she responds to the reasons that it seems to her there are for forming that belief, not the existence of such reasons. And, on a certain crucial assumption, a subject's having the desires that she is required to have by the norms of practical rationality entails neither the existence of reasons for forming that desire nor that it seems to her that there are such reasons for forming that desire. Unlike theoretical rationality, practical rationality is not even a matter of being responsive to what seem to be reasons.

However, to repeat, this is so only on a certain crucial assumption, namely, that the norms of practical rationality are all like the means-ends norm in being requirements to have certain desires, given that we have certain other desires. There are two main strategies that can be pursued in the attempt to reject this assumption: the Besire

Strategy and the Rationalist Strategy. The Besire Strategy requires us to find values of 'p' and 'φ' that make a principle of the following form come out true:

BD: If an agent believes that p then she desires to φ.

The Rationalist Strategy requires us to find values of 'p' and 'φ' that make a principle of the following form come out true:

PR<sup>R</sup>: Reason requires that if an agent believes that p, then she desires to φ.

Each strategy, if successfully pursued, would provide us with an account of how there could be reasons for desires in the sense of considerations that justify our having certain desires.

However, I have also explained why, as it seems to me, the Besire Strategy is not very plausible, and why we are bound to have at least moderate success when we pursue the Rationalist Strategy. There is at least one consideration that justifies desire, namely, the fact that the desirer herself would want herself to act in the way desired if she had a desire set that conforms to all of the norms of practical reason. Whether there are further considerations that justify desires is an open question, one which must be pursued on another occasion.

#### REFERENCES

Dancy, Jonathan 2000: Practical Reality (Oxford: Oxford University Press).

Davidson, Donald 1963: 'Actions, Reasons and Causes' reprinted in his *Essays on Actions and Events* (Oxford: Oxford University Press, 1980), pp.3-20.

Hume, David 1740: A Treatise of Human Nature (Oxford: Clarendon Press, 1968).

McDowell, John 1978: 'Are Moral Requirements Hypothetical Imperatives?', Proceedings of the Aristotelian Society Supplementary Volume 52, pp.13-29.

Nagel, Thomas 1970: The Possibility of Altruism (Princeton: Princeton University Press).

Parfit, Derek 1997: 'Reasons and Motivation' in Proceedings of the Aristotelian Society Supplementary Volume 71, pp.99-130.

Pettit, Philip and Michael Smith forthcoming: 'External Reasons' forthcoming in McDowell and His Critics edited by Cynthia Macdonald and Graham Macdonald (Oxford: Blackwell).

Geoffrey Sayre-McCord, 'The Meta-Ethical Problem: a discussion of Michael Smith's The Moral Problem' in Ethics, 108, 1997, pp.55-83.

Smith, Michael 1994: The Moral Problem (Oxford: Blackwell Publishers).

\_\_\_\_\_ 1997: 'In Defence of The Moral Problem: A Reply to Brink, Copp and Sayre-McCord' in Ethics, 108, pp.84-119.

\_\_\_\_\_ forthcoming: 'Instrumental Desires, Instrumental Rationality' in Proceedings of the Aristotelian Society Supplementary Volume.

Wallace, R.Jay 1990: 'How to Argue about Practical Reason' in Mind, 99, pp.267-97

Williams, Bernard 1980: 'Internal and External Reasons' reprinted in his Moral Luck (Cambridge: Cambridge University Press, 1981).

Woods, Michael 1972: 'Reasons for Action and Desire', Proceedings of the Aristotelian Society Supplementary Volume 46, pp.189-201.