

XIV*—REASON AND DESIRE

by Michael Smith

I

My topic is the debate in moral psychology between the rationalist and the anti-rationalist over the proper relation between reason and desire. My aim is not to adjudicate this debate, but rather to clarify what is at stake, for, it seems to me, both parties are prone to misconceive the issues that divide them.

Rationalists and anti-rationalists alike believe that there are norms of rationality that govern the way we come to desire what we desire and do what we do, much as there are norms of rationality that govern the way we come to believe what we believe. This is as it should be. For claims that give characteristic expression to such norms play an integral role in our everyday lives, claims like '*A* ought to believe that *p*', '*A* has a reason to believe that *p*', '*A* ought to want to Φ ', '*A* has a reason to want to Φ ', '*A* ought to Φ ' and '*A* has a reason to Φ '.¹ Disagreement begins not with the claim that there are such norms, but with the rationalist's and anti-rationalist's different theories about the content of such norms and the relations between them. For rationalists claim, first, that rationality places certain practical requirements on action, requirements that are not reducible to requirements of theoretical reason, and second, that the requirements of morality reduce to these practical requirements of reason. Anti-rationalists, for their part, simply deny this conjunction.²

We seem to have here a fairly straightforward debate. But can

* Meeting of the Aristotelian Society held at 5/7 Tavistock Place, London WC1, on Monday, 13 June, 1988 at 6.00 p.m.

¹ The reasons mentioned here are of course all *normative* reasons, as distinct from *motivating* reasons. As I understand it, normative reasons simply are norms governing an agent's beliefs, desires or actions. For more on this distinction see my 'The Humean Theory of Motivation' in *Mind* January 1987, pp. 37-41.

² 'Anti-rationalist' is a term of art. Anti-rationalists may certainly find plenty of room for rational considerations in arguing for their moral views. Consider, for example, R. M. Hare's 'Ethical Theory and Utilitarianism' and Thomas Scanlon's 'Contractualism and Utilitarianism' both in Amartya Sen and Bernard Williams' *Utilitarianism and Beyond* (Cambridge, 1982). They simply deny that rationality can play the *foundational* role it plays in a theory like Kant's or, more recently, Thomas Nagel's (see his *The Possibility of Altruism* (Princeton, 1970)).

we narrow it further? Or is there more at stake than at first appears?

II

The most famous of the anti-rationalists, Hume, denies *both* the rationalist's claims. But why does he deny the first, the claim that there are irreducible norms of practical reason?

Hume thinks that when we call some action or desire unreasonable or irrational we are simply criticising it on the basis of the beliefs from which it was derived.³ The central cases are thus supposed to be like the following:⁴ someone wants to drink a gin and tonic and drinks from the glass before him believing it to contain gin and tonic when in fact it contains gin mixed with petrol. We certainly find ourselves saying that he wants to drink from the glass even though he *has no reason* to do so, in such cases, or that he *shouldn't* want to do so. And these normative claims seem to be made from, in the broad sense, the perspective of rationality, or reason.⁵ But the explanation, according to Hume, is that the man's wanting to drink from the glass and his doing so both depend crucially upon his having the belief that the glass contains gin and tonic, a belief that fails in terms of what is desirable in the way of a belief, namely, *truth*. Thus, he says, our claim that such a man has a desire that he shouldn't have can be seen to depend on the fact that that desire is derived from, *inter alia*, a belief that he shouldn't have, and thus, ultimately, on the norms of theoretical reason.

This treatment suggests the following quite general reductive strategy. We begin by formulating some norms of theoretical reason like the following:⁶

If *p*, then *A* believes that *p*, and if not-*p*, then *A* believes that not-*p*.

We then simply list all the ways in which a man may be criticised

³ See his *Treatise of Human Nature* edited by Selby-Bigge (Oxford, 1888) pp. 413–7. The argument that follows owes a great deal to Bernard Williams' 'Internal and External Reasons' in his *Moral Luck* (Cambridge, 1981) and Christine Korsgaard's 'Skepticism about Practical Reason' in *Journal of Philosophy* January 1986 (especially section III).

⁴ The example comes from Williams' 'Internal and External Reasons' p. 102.

⁵ For more on this 'broad sense' see footnote 9 below.

⁶ This principle comes from Bernard Williams' *Descartes: The Project of Pure Inquiry* (Pelican Books, 1978) p. 38.

for his desires on the basis of such norms, hoping that our list is exhaustive of the ways in which desires may be subjected to rational criticism. Examples of Hume's kind depend on failures of theoretical reason characterisable in terms of the second half of this norm. Moreover, as we might hope, there are examples that depend on failures characterisable in terms of the first half. Suppose someone wants to buy a Picasso, and, unbeknown to him, the painting before him is a Picasso. We would certainly find ourselves saying that he *has a reason* to buy the painting before him, or that, though he doesn't want to buy it, he *should*. But, as before, we can say that the normative force of these claims depends on what is desirable in the way of beliefs: in this case, on the fact that the agent fails to believe one of the truths, and that he should therefore have had the desire that he could have derived from the belief he should have had together with his desire to buy a Picasso.

However, as Kantians are fond of pointing out, this strategy proves to be inadequate. For there are kinds of failure of practical reason inexplicable in anything like these terms. Indeed, the idea that the strategy succeeds even in the cases we have so far considered depends on ignoring one potential failure, for it takes for granted the idea of 'deriving' a further desire from a background desire and belief.⁷ To see this it is useful to consider a kind of failure of theoretical reason unlike those already mentioned. Suppose A believes both that p and that $p \rightarrow q$, but does not believe that q . We certainly find ourselves saying that A *has a reason* to believe that q , or that he *ought* to have that belief. Indeed, we find ourselves saying this even when q is false.⁸ But why? The norm mentioned above doesn't tell us. The

⁷ In what follows I thus restrict my discussion to failures of means-end rationality, the example most often discussed by Kantians (see for example, *The Possibility of Altruism* pp. 33-4), for the Kantian response to this example proves to be important in what follows (sections III and IV). There are, however, doubtless other failures of rationality that Hume's strategy cannot handle—genuine weakness of will, for example.

⁸ The 'oughts' here and elsewhere in this paper, and thus the normative reason claims as well, must all be understood as being merely *prima facie*. Only so can we explain why we can coherently hold that an agent ought to believe that q and ought to believe that not- q . The problem vanishes once we realise that he ought to believe not- q relative to the fact that not- q , and q relative to the fact that he believes both that p and that $p \rightarrow q$. A similar problem might be thought to arise in the desire case where the same stipulation removes the apparent contradiction. For the treatment of '*prima facie*' that secures this result see Donald Davidson's 'How is Weakness of the Will Possible?' in his *Essays on Actions and Events* (Oxford, 1980) p. 38.

answer lies in the fact that a subject of beliefs is capable of reflection in accordance with principles allowing us to derive truths from truths. Such reflection must therefore itself be regulated, albeit indirectly, by the aim of truth. It thus follows that there are additional norms of theoretical reason, in particular, a norm telling us that we ought to be disposed to form new beliefs on the basis of old in accordance with *modus ponens*.⁹ Now the charge against Hume can be put like this: he fails to see that certain failures of reason require explanation in terms of analogous norms of *practical* reason specifying what is desirable in the way of dispositions to form new desires on the basis of old desires and beliefs.¹⁰

Suppose *A* desires to buy a Picasso and believes that the painting before him is a Picasso, but that *A* does not desire to buy the painting before him. That is, suppose *A* fails to desire the believed means to one of his ends. This would seem to involve *A* in irrationality of a fairly gross kind. But can we characterise the kind of irrationality in terms of a failure of *theoretical* reason? We cannot. For we can imagine both *A*'s beliefs and his dispositions to form new beliefs on the basis of old being just as they should be. This kind of failure seems rather to require explanation in terms of a norm expressing what is desirable in the way of *desires*, a norm like the following:

If *A* desires that he Φ 's and he can Φ by Ψ -ing then *A* desires that he Ψ 's

For just as the considerations adduced earlier allow us to see how

⁹ It seems that violations of this kind of norm of theoretical reason alone make appropriate in ordinary language the charge of *irrationality*. In ordinary language we are apparently more concerned with failures of theoretical *reasoning* than with the more general subject of failures of theoretical *reason* (thus the 'broad sense' of the perspective of rationality or reason mentioned above). Korsgaard herself seems to me to be taken in by this ordinary language point in her 'Skepticism about Practical Reason' pp. 11–12.

¹⁰ In what follows I assume, for the sake of simplicity, that an account of what is desirable in the way of *beliefs* is the concern of norms of theoretical reason and that an account of what is desirable in the way of *desires* is the concern of norms of practical reason. I put to one side the problem of explaining how beliefs about the desirability of performing certain actions, or moral beliefs, which are clearly governed by norms of theoretical reason insofar as they aim to be *true*, must nonetheless be governed in some way by norms of practical reason too, given that they have *practical* implications. My views on this more complicated matter depend on arguments to be found in section 5 of 'The Humean Theory of Motivation'. Nothing said here turns on these arguments, however.

'If A believes that p and he believes that $p \rightarrow q$ then he ought to believe that q ' expresses a norm governing dispositions to form new beliefs on the basis of old beliefs, so the principle just formulated allows us to see how 'If A desires that he Φ 's and he believes that he can Φ by Ψ -ing then he ought to desire that he Ψ 's' expresses a norm governing dispositions to form new desires on the basis of old desires and beliefs. Just as the former expresses an irreducible norm of *theoretical* reason, so the latter expresses an irreducible norm of *practical* reason. Thus (the argument goes) Hume is wrong to suppose that we can reduce the requirements of practical reason to the requirements of theoretical reason.¹¹

This strikes me as a perfectly good argument. However, granting this point doesn't give us any reason to decide the debate between the rationalist and the anti-rationalist. For the anti-rationalist can agree that there are irreducible norms of practical reason, like that concerning means-end reasoning just mentioned, while still insisting that these norms do not provide a reductive base for the norms of morality. For moral norms concern what an agent ought to do or want *simpliciter*, not just what he ought to do or want *given* that he already wants other things. From the perspective of the most plausible form of anti-rationalism, then, Hume and Kant are both extremists: one's claims about the norms of practical reason are too austere while the other's claims are too extravagant.

III

The debate can thus be narrowed. Anti-rationalists should accept the rationalist's first claim and focus their disagreement on the second. However there now emerges a potential source of deeper division. For if the success of rationalism depends on the rationalist's being able to show that there are norms of reason

¹¹ These practical principles would require more careful formulation in a fuller treatment. Suppose I want money (as I do) and believe (as I do) that I could get it by killing my wife and collecting the insurance. Am I irrational if I do not desire to kill my wife? Surely not! What this shows is that certain courses of action that would get me what I want are, *given other things that I want*, off the agenda as means to my ends. A more careful formulation would thus show a more complicated dependence of an agent's desires for the means on his desires for the end in the context of the rest of his desires. For more on this feature of desires for means see Bernard Williams' 'Practical Necessity' in his *Moral Luck*.

telling us what agents ought to do, or want, *simpliciter*, not just what they ought to do, or want, *given* that they want certain other things, then, in arguing for his second claim, the rationalist may seem to commit himself to the view that *reason* can be a motive, either a motive directly for the actions we ought to do, or for the desires we ought to have.

Rationalists certainly seem to accept this implication, for they spend a good deal of time defending the claim that reason can be a motive.¹² And at least some anti-rationalists seem to accept it too, for they try to argue against rationalism by appealing to the claim that reason cannot be a motive.¹³ Rationalists, in their turn, have recently argued that this counter-argument fails; that no argument could establish that reason cannot be a motive except an argument directly against rationalism itself.¹⁴ I find myself in disagreement with much of this. Let me begin by explaining why rationalists and anti-rationalists can and should agree that reason cannot be a motive. I will concentrate on Christine Korsgaard's recent argument to the contrary.

Korsgaard claims that there are two sorts of doubt one might have about the extent to which action is, or could be, directed by reason (p.5). The first is doubt about whether there are any rational principles that give substantive guidance to choice and action—about whether there are any irreducible requirements of practical reason. This she calls *content skepticism*. The second is doubt about 'the scope of reason as a motive'—about whether any belief could serve as a motive. This she calls *motivational skepticism*. In these terms, Korsgaard's aim is to show that motivational skepticism is wholly dependent on content skepticism. She argues as follows.

Blanket content skepticism, doubt about whether there are *any* irreducible norms of practical reason, is Hume's view. But, for reasons already given, Hume is wrong: the means-end principle expresses an irreducible norm of practical reason. Of course, accepting this is consistent with denying that beliefs can be motives. For, as Korsgaard notes, when rational persons are motivated in accordance with the norms governing means-end

¹² See Nagel's *The Possibility of Altruism*.

¹³ See Williams' 'Internal and External Reasons'.

¹⁴ See Korsgaard's 'Skepticism about Practical Reason'.

reasoning the desire for the means has a 'clear motivational source' in the desire for the end (p.11). However, she insists, it does suffice to show that we have no reason to rule out *without further argument* the possibility of *further* principles expressing norms of practical reason, principles connecting beliefs with desires directly—principles expressing what's desirable in the way of desires like the following:¹⁵

If someone desires that p and A can promote p by Ψ -ing then A desires that he Ψ 's.

But if this were to express an irreducible norm of practical reason then, by analogy with the belief case, 'If A believes both that he can promote p by Ψ -ing and that someone desires that p then A ought to desire that he Ψ 's' would express a norm governing dispositions to form desires on the basis of beliefs. And since this is equivalent to 'If A believes both that he can promote p by Ψ -ing and that someone desires that p then, if A is rational, A desires that he Ψ 's', so, again by analogy with the belief case, A 's belief that he can promote p by Ψ -ing together with his belief that someone desires that p would, if A were rational, give rise in him to the desire to Ψ (just as, in a rational person, the beliefs that p and that $p \rightarrow q$ give rise in him to the belief that q). To repeat, nothing so far said justifies us in thinking that there *is* such a norm. Korsgaard's point is that nothing so far said justifies us in thinking there *isn't*—that what needs to be addressed is what is and what is not an irreducible norm of practical reason. And that, she claims, suffices to show that motivational skepticism *depends* on a kind of content skepticism. For if, for all we know, there is such a norm, then, for all we know, A 's believing that he can promote p by Ψ -ing and that someone desires that p may be a belief that functions as a motive—a motive for the desire to Ψ and any action that results from it.

Should we accept this argument? I think not. For the time being let us assume that rationalism requires that there are norms like the ones just mentioned. (I take back the concession

¹⁵ The following principle is meant to express the kind of norm Nagel argues for in Part 3 of *The Possibility of Altruism*. If we were considering the kinds of norm that some other rationalist, say, Kant, argues for, then we would need to make corresponding adjustments in the argument that follows, especially in the next section. I will simply assume that the adjustments are themselves fairly transparent.

in the next section.) Even if there were such norms, we should deny that the beliefs that 'give rise to' desires in rational persons serve as their *motives*. To see why, compare the two principles that are said to be analogous: 'If A believes that p and he believes that $p \rightarrow q$ then A ought to believe that q ' and 'If A believes that he can promote p by Ψ -ing and that someone desires that p then A ought to desire that he Ψ 's'. Suppose A acquires both the belief that q and the desire that he Ψ 's in accordance with these norms. Korsgaard assumes that it would be appropriate to hold that, in the belief/desire case, the antecedent belief *motivates* the desire. But would that be appropriate in the belief/belief case? When I come to believe that q because I believe both that p and that $p \rightarrow q$, are these beliefs my *motives* for believing that q ? Surely not. (Indeed, in ordinary language a motivated belief is precisely one that is not susceptible to normal rational explanation in terms of other beliefs. It is rather a belief held because it is desired, as in wishful thinking.) But if it is wrong to say that the belief motivates the belief in the belief/belief case, then shouldn't it be wrong to say that the belief motivates the desire in the belief/desire case?

I think that it is wrong in both cases, and that it is not surprising that it is wrong. For there are conceptual links between the concept of a *motive* and the concept of a *goal*. In short, a motive is a psychological state that embodies the having of a goal. For beliefs to serve as motives, therefore, a belief would have to be able to embody the having of a goal. But that is not possible. For, for these purposes, the crucial feature of a belief is what has been called its *direction of fit*.¹⁶ Beliefs are states with a content that is supposed to fit the world. Finding that the content of a belief does not fit the way the world is, is ground for changing the belief. But a state that embodies the having of a goal has just the opposite direction of fit. It has a content that the world is supposed to fit. Thus finding that the world does not fit the content of a state that embodies the having of a goal is ground for changing the world, not for changing the state. Beliefs are thus *in their nature* unsuited to embodying the having of goals. For the only kind of state that could do that is a state that, unlike

¹⁶ See G. E. M. Anscombe's *Intention* (Blackwell, 1957) section 2 and Mark Platts' *Ways of Meaning* (Routledge and Kegan Paul, 1979) pp. 256-7.

a belief, has the same direction of fit as a state that embodies the having of a goal. The motivational skeptic insists that *desires* are the only psychological states with a content that the world is supposed to fit. And, as I have argued elsewhere, this seems plausible.¹⁷

If this is right, the rationalist himself should be a motivational skeptic. Korsgaard is wrong that the only argument for motivational skepticism is a direct argument to the effect that there are no norms of the belief/desire kind, as are anti-rationalists who hold that motivational skepticism gives them reason to doubt that there are such norms. Motivational skepticism, itself neutral about the existence of such norms, rather gives both theorists reason to insist that, if there were such norms then, if *A* were rational, his believing that he can promote *p* by Ψ -ing and that someone desires that *p* would *give rise to*, and thus *explain*, his desiring to Ψ without *motivating* it. For, if the rationalist is right, there are explanatory connections between beliefs and desires of *the same kind* as those we find elsewhere merely between beliefs, a species of rational explanation that is, unlike explanations involving motives, *non-teleological*. The point is not merely terminological. The rationalist misrepresents the *substance* of his own view if he insists, as Korsgaard does, that beliefs may be motives.

Nor is Korsgaard alone in misrepresenting rationalism in this way. In *The Possibility of Altruism* Nagel insists that though a desire underlies every intentional action, that is so only because some of these desires are themselves *motivated* desires, desires 'arrived at by decision and after deliberation' (p.29). Though he agrees that certain desires are motivated by other desires, he claims that the issue behind rationalism is whether 'another desire always lies behind the motivated one, or whether sometimes the motivation of the initial desire involves no reference to another, unmotivated desire', but rather a belief (p.29). But if that were the issue it would be easily resolved. For if a desire is motivated then it is susceptible to teleological explanation. The motive that motivates the desire must therefore embody some goal of the agent. But no belief could do that, for beliefs are not the right kind of state to embody the

¹⁷ See section 5 of my 'The Humean Theory of Motivation'.

having of goals. The conclusion is not that rationalism is untenable. It is rather that Nagel misrepresents the issue. The issue may well be whether beliefs can rationally explain desires. But, if it is, it has nothing to do with whether beliefs can be motives.¹⁸

IV

Rationalists might regroup. They might suggest a new source of deeper division. For it seems that the rationalist, unlike the anti-rationalist, is committed to holding that reason can *give rise to or produce* a motive in much the way that beliefs produce beliefs in rational inference.¹⁹ I want now to argue that this too is misleading. Though one kind of rationalist, the *belief-rationalist*, does indeed claim that reason may produce a motive, another, the *desire-rationalist*, rejects that claim.

What is the difference between these two kinds of rationalist? Let us suppose both agree that all fully rational beings desire to Ψ when they believe that someone desires that p and that they can promote p by Ψ -ing. The difference between them lies in their *explanations* of this fact. As we have seen, the belief-rationalist thinks that it is simply part of the ideal rational function of pairs of the relevant beliefs to produce the relevant desires, much as it is part of the ideal rational function of the belief that p to produce the belief that q in conjunction with the belief that $p \rightarrow q$. The desire-rationalist rejects this. He thinks instead that the beliefs that the belief-rationalist thinks produce the desire to Ψ all by themselves, are just what they seem to be: means-end beliefs that can produce a desire only in conjunction with a background desire. He thinks that the reason all rational creatures desire to Ψ in the presence of such beliefs is that they all desire to promote the satisfaction of each person's desires, and that they therefore derive the desire to Ψ in the normal means-end way. What makes the desire-rationalist's view distinctively rationalist is that, according to him, the norms of reason tell us that one of the things that is desirable in the way of desires is precisely

¹⁸ For a discussion of this passage from *The Possibility of Altruism* from a slightly different vantage point see my 'The Humean Theory of Motivation' pp. 58–60.

¹⁹ Jay Wallace convinced me of the importance of considering this claim on its own terms. Wallace's views may be found in his *Practical Reason and the Claims of Morality: the Idea of Rationalism in Ethics*, PhD dissertation, Princeton university 1988.

*A's having the desire to promote the satisfaction of each person's desires.*²⁰

The belief- and desire-rationalist thus offer us two quite different ways of giving expression to rationalism. Is one of these preferable? I think not. Most rationalists will, I suspect, disagree. They will think that belief-rationalism is better. But they will be hard pressed to say why. Perhaps they will argue that the belief-rationalist can explain failures of practical reason that the desire rationalist cannot. Let us consider that claim.

The belief-rationalist holds that my failure to desire to Ψ when someone desires that p and I can promote p by Ψ -ing may be explained in several ways. One is that I do not believe either that someone desires that p , or that I can promote p by Ψ -ing. My failure is then ignorance.²¹ But the desire-rationalist agrees that one explanation of my failure will appeal to ignorance, for he agrees that when I have a certain background desire I may fail to derive further desires in accordance with the norms governing means-end reasoning if I fail to believe that the means is a means to my end (see section II). And if I have the desire to promote the satisfaction of each person's desires, then, when someone desires that p , failing to believe that I can promote p by Ψ -ing, or that someone desires that p , are precisely ways of being ignorant about the means to my end.

The belief-rationalist thinks that we can explain my failure to desire to Ψ even when I am not ignorant. For he thinks that the connection between belief and desire is akin to the connection between belief and belief in rational inference. He thus claims that I might fail to desire to Ψ because, though I would desire to Ψ if I were fully rational, something may prevent me from being fully rational, so preventing me from making the relevant rational connection. Korsgaard puts the point in this way: 'Rage, passion, depression, distraction, grief, physical or mental illness: all these things could cause us . . . to fail to be motivationally responsive to the rational considerations available to us' (p.13), just as (she might have added) they may be

²⁰ Though he does not endorse this view, Derek Parfit clearly recognises its coherence in Chapter 6 of his *Reasons and Persons* (Oxford, 1984).

²¹ According to the belief-rationalist I am like the man who, despite the fact that p , that $p \rightarrow q$ and hence that q , fails to believe that q because (in addition to not believing q) he believes neither that p nor that $p \rightarrow q$; or because, though he believes p , he doesn't believe that $p \rightarrow q$; or because, though he believes that $p \rightarrow q$, he doesn't believe that p .

responsible for our irrationality when we fail to perform *modus ponens*.

However, the desire-rationalist thinks there is an explanation of my failure like this as well. Indeed, he thinks that there are two such explanations. First, since he holds only that all *rational* creatures have the desire to promote the satisfaction of each person's desires, he thinks that, if I am not fully rational, perhaps because I am enraged, impassioned, depressed, distracted, etc., then I may not have that desire. But then even if I do believe that I can promote p by Ψ -ing and that someone desires that p , I will not desire to Ψ . For, according to him, I will have that desire only if I have a background desire to motivate it. Second, even if I have both the background desire to promote the satisfaction of each person's desires, and the relevant beliefs, I may still not desire to Ψ because I may fail to conform my desires to the means-end principle (section II)—that is, because I fail to derive a desire for the believed means to my end. Here too, he will say, we may explain such failures of reasoning in terms of factors like rage or distraction, factors that may prevent me from being fully rational. Indeed these are, he will say, the real analogues of the failures of theoretical reasoning that the belief-rationalist appeals to; these are the cases in which I really fail to make the 'relevant rational connection'. It thus seems that both belief- and desire-rationalism will furnish the rationalist with explanations of the failures of practical reason that he seeks.²²

²² Is there another kind of explanation that both the belief- and desire-rationalist could give of my failure to desire to Ψ when I believe both that I can promote p by Ψ -ing and that someone desires that p ? It might be thought that there is. After all, in belief-rationalist terms, mightn't I simply fail *altogether* to have the disposition to make the rational connection between these beliefs and the desire to Ψ ; that is, fail despite the fact that I am not enraged or impassioned or. . . ? Mightn't I be simply *incapable* of making the relevant rational connection? Or, in desire-rationalist terms, mightn't I simply fail to desire to promote the satisfaction of each person's desires despite the fact that I am not enraged. . . and so on? Mightn't I be simply *incapable* of having a desire that is rationally required? Of course, we would need then to *ground* the charge of irrationality. But perhaps the idea could be that we have a *social* conception of (practical) rationality. The issue is delicate, for rationalists can agree to this only if they deny a claim that many rationalists and anti-rationalists, theorists as diverse as Kant and Hume, Nagel and Williams, seem simply to have taken for granted: the claim that practical reasons of rationality must satisfy what Korsgaard calls the *internalist requirement* ('Skepticism about Practical Reason' pp. 11-5). As Williams construes that requirement, an agent has a normative reason of rationality to Φ (that is, he is rationally required to Φ) only if, if he were to do some kind of 'deliberative reasoning', he would desire to Φ ('Internal and

How might the belief-rationalist argue that his way of giving expression to the rationalist's view is preferable? He might try arguing that, though the desire-rationalist can give *surrogates* of each of the explanations he gives of failures of practical reason, his explanations are *better* because they are more *intuitively compelling*. But why is this so? The only reason could be that our *ordinary* explanations of desire acquisition somehow commit us to the view that beliefs can by themselves rationally explain desires. However, this claim seems altogether implausible.

We can, of course, think of contexts in which nothing over and above such beliefs need be mentioned in order to make the acquisition of such a desire seem appropriate. Consider Nagel's example of coming to desire to get off a man's gouty toe having realised that it is one's own heel that is causing the pain.²³ But is it really obvious that in such cases we don't, as Davidson has suggested, simply take the background desire for granted?²⁴ However I don't want to insist on that. For I also find implausible the counter-claim that our ordinary explanations

External Reasons' pp. 102–6). But if there is to be an explanation of my failure like that mentioned above, then we must deny this. For the explanation undermines the *possibility* of my deliberating to a desire to Ψ . Korsgaard's formulation of the requirement makes the problem even more explicit, for she claims that an agent is rationally required to Φ only if, if she knew the considerations that were supposed to rationalize the desire (that is, if she had the beliefs that the belief-rationalist thinks rationally explain the desire), and 'if nothing were interfering with her rationality', then she would desire to Φ . She goes on to insist that this isn't trivial 'as long as the notion of a psychological condition that interferes with rationality is not trivially defined' (p. 14), which merely takes us back to her list: rage, passion and the like. But the kind of failure that we are contemplating here is precisely one in which these *cannot* explain my failure. Should rationalists reject the possibility of such an explanation or reject the internalist requirement? I am not sure. Interestingly, the best argument I know for rejecting the requirement comes from Korsgaard herself. (This is ironic, given that she accepts the requirement.) For she an example that seems to indicate that we can imagine someone who is just *incapable* of coming to desire the believed means to his ends: i.e. someone who fails to do so despite the fact that he is not enraged, impassioned, depressed, distracted, grieved, or physically or mentally ill. Yet, if the means-end principle is a norm of reason at all, such a person clearly *has a reason* to do so. For those who doubt that we can imagine this, as I once did, consider the passage from Kant that Korsgaard cites with approval in which Kant himself describes for us, in Korsgaard's words, 'a human being in . . . a condition of being able to reason, so to speak, theoretically, but not practically' (p. 13). It goes without saying that, if we were to reject the internalist requirement, the method of arguing for or against rationalism would seem to be entirely up for grabs.

²³ See *The Possibility of Altruism* pp. 84–5.

²⁴ See Donald Davidson's 'Actions, Reasons and Causes' in his *Essays of Actions and Events* pp. 6–7.

somehow commit us to the view that the only rational explanation of a desire is a means-end explanation. The problem here is that our ordinary talk is simply too unsystematic to commit us *either* to the view that beliefs may by themselves explain desires *or* to the opposite view. Ordinary talk leaves indeterminate the details of just the examples we need to appeal to, and filling in the details requires that we interpret the examples in the light of a theory giving us a systematic philosophical account of ordinary psychological states. What the belief- and desire-rationalist are trying to decide is precisely what the best such theory would be like. Would the theory be constructed in terms of the norms of belief- or desire-rationalism? Thus, though we should agree that our ordinary explanations certainly *constrain* our choice of theory, we should, I think, reject the belief-rationalist's claim that an appeal to ordinary explanations tells especially in favour of his theory. As far as the intuitiveness or otherwise of the different theorists' explanations go, both theories seem equally acceptable.²⁵

If this is agreed, then it seems that we have little left to choose between belief- and desire-rationalism.²⁶ That has significant implications for what we, at present, take to be the issue between the rationalist and the anti-rationalist. For the desire-rationalist agrees with the anti-rationalist that the production of certain desires is no part of the ideal rational function of any belief. Like the anti-rationalist he thinks that the beliefs that the belief-rationalist claims can produce certain desires all by themselves are simply means-end beliefs that can produce desires only in conjunction with some appropriate background desire. But he differs from the anti-rationalist in holding that certain background desires are not a wholly contingent feature of a rational

²⁵ By which, of course, I mean to leave open the possibility that both theories are unacceptable because anti-rationalism is correct.

²⁶ This is not to say that we mightn't ultimately find reason to prefer one or the other theory—say, because one better coheres with *other* theories, for example, with the best theory of practical reasoning. It is too early to adjudicate this issue, for discussions of practical reasoning to date remain programmatic and inconclusive. For a discussion of belief-rationalist theories of practical reasoning see Parts 2 and 3 of Nagel's *The Possibility of Altruism* and Chapters 8–10 of his *The View From Nowhere* (Oxford UP, 1986). Nagel offers some criticisms of desire-based theories of practical reasoning in Chapters 5–8 of *The Possibility of Altruism*. For an attempt to respond to at least some of these criticisms see Derek Parfit's discussion of the Present Aim Theory in Part 2 of his *Reasons and Persons*.

creature's psychology. He thinks, instead, that certain desires, desires like the desire to promote the satisfaction of each person's desires, are desires that fully rational creatures have *necessarily*. The belief-rationalist, by contrast, agrees with the anti-rationalist that the desire-rationalist is wrong to suppose that certain desires are categorically required by reason. But he differs from the anti-rationalist in holding that part of the ideal rational function of certain beliefs is the production of certain desires. It is thus, at the very least, misleading for rationalists to suggest that a source of deep division between themselves and the anti-rationalists is whether reason can produce a motive. For that is so on only one interpretation of what rationalism is.

V

I began this paper saying that I wanted to get clear what is at stake in the debate between the rationalist and the anti-rationalist. The upshot seems to be that there is, at present, no single issue that divides them over and above whether moral norms reduce to the norms of practical reason. For, we have seen, rationalists and anti-rationalists should both agree that there are irreducible requirements of practical reason (as against Hume), and that reason cannot be a motive (as against Nagel and Korsgaard). Beyond this, we have seen that rationalists may disagree among themselves over the virtues of belief- and desire-rationalism, a disagreement that has significant effects on what they take to be at stake in their debate with the anti-rationalist. Eventually, of course, we may find reason to prefer belief-rationalism, or to prefer desire-rationalism, in which case a further source of deep division will emerge. But at present the issue between them is simply indeterminate.

In the face of such indeterminacy, how should rationalists proceed? It seems to me that, barring further considerations that decide the issue, rationalists should simply stipulate how they intend to formulate their claims, whether that be in belief- or desire-rationalist terms, and then proceed to make their claims in those terms. But, if the foregoing discussion is right, it suggests that, having chosen to formulate their claims in these terms, the rationalist must not make the mistake of thinking that there *must* be some further deep issue about whether reason can or cannot (depending on how he chooses to formulate his claims), produce

a motive, or whether there are or are not (again depending on how he chooses to formulate his claims) desires that are categorically required by reason. For he could suppose that only if he mistakenly thinks that a potentially arbitrary formulation of his claims is itself required by reason.²⁷

²⁷ I would like to thank Simon Blackburn, Stewart Cohen, Philippa Foot, Robert Gay, Barbara Herman, Brad Hooker, Mark Johnston, Christine Korsgaard, Philip Pettit, Gideon Rosen and, especially, Jay Wallace, for their many helpful comments.