Reply to Bratman and Smith
KIERAN SETIYA

To begin with, I am deeply grateful to Michael Bratman and Michael Smith for their generosity in responding to my book, for the care with which they have read it, and for the challenge of meeting their objections. I am also grateful for their support and encouragement over the years. It is a pleasure to engage with them here.

Because their comments raise many related difficulties, this reply will treat them together, beginning with brief consideration of issues in action theory before turning to Reasons and rationalism. It will also be incomplete: there is much more to say about these problems than can be said in this space.

Bratman’s discussion poses a structural question: how far does my argument about reason and virtue in Part Two of the book depend on the principle of Belief and the cognitive theory of intention from Part One? His answer is: quite a bit. But this is not straightforward. If we reject Belief, our account of acting for reasons becomes more minimal. This would make it more difficult to derive standards of practical reason from the nature of agency, as the rationalist purports to do. If I am right to argue that rationalism and the virtue theory are exhaustive alternatives, it therefore makes the virtue theory of practical reason easier to defend. My Anscombean assumptions are in fact a concession to some of the ideas that motivate rationalists like Korsgaard and Velleman – though not, as we will see, either Bratman or Smith. Denying them makes more trouble for the ethical views that I oppose than it does for mine.1

1 Still, I do hope to defend Belief, or something like it, as a pivotal constraint in action theory. Bratman makes a series of objections, of which I briefly respond to three. First, he thinks that there are cases in which I am doing φ intentionally without the belief that I am doing it or the belief that I am doing anything else as a means to that end – except for trying to φ. Even if he is right about this, however, it might be sufficient for my arguments. We can still ask why the belief that I am trying to φ must be present. (A different response, in Setiya 2008: 390–2, leans on the possibility of partial belief.) Second, even if Belief is true of intentional action, Bratman gives plausible cases of prospective intention in which I do not believe that I will act accordingly, and may not even be sure that I will try. Here, I am willing to be liberal about the word ‘intend’, which may apply in the cases that Bratman describes. What matters is that beliefs about what one is doing are essential in acting for reasons, and that, cast in the prospective mode, they form the basis of planning agency. Finally, Bratman doubts the implication of my view that one cannot have an intention one does not believe one has. Questions about self-knowledge are too large and too difficult to be examined here; they now seem to me absolutely crucial. In fact, the best argument for a cognitive theory of intention may draw on considerations of ‘transparency’ more familiar in connection with belief. I pursue this line of argument in ‘Knowledge of Intention’ (Setiya, unpublished).
What does matter to my argument is that when we act intentionally, or for reasons, we need not do so ‘under the guise of the good’, since the doctrine that we must might well support a form of ethical rationalism. Smith is not convinced that I establish this crucial claim. He asks what is to prevent my opponent from urging that the desire-like or motivating belief involved in taking something as one’s reason, according to me, constitutes an appearance of the good? How do I rule out this conception of desire, if not by the counterexamples I admit to be inconclusive?

In answering these questions, it is useful to say more about the methodology of my approach to intentional action. It rests on the view that metaphysical necessities must be explained by the essence or nature of their objects. In the simplest and central case, the necessary truth that all Fs are G is explained by a metaphysical definition of what it is to be F partly in terms of being G. It is this demand for explanations of necessity that lies behind the puzzle from Anscombe that Smith describes. Roughly put, why must doing something for reasons involve the belief that one is doing it? How is the truth of Belief entailed by the nature of intentional action? Now, those who defend the guise of the good accept a necessity of their own:

If A is doing φ intentionally, or for reasons, φ-ing must appear to him in some way good.

In particular, I focus on the following proposal:

If A is doing φ on the ground that p, he must believe that the fact that p is a normative practical reason for him to φ.

How do I argue against this claim? By giving an account of what it is to act on the ground that p that does not require that normative thought, but which explains everything else one needs to explain with such an account, as for instance the truth of Belief. If this account is right, thinking that one’s reason for doing φ is a normative reason is an optional extra.

Smith’s concern is that, even if I have refuted one version of the guise of the good, another might be hidden in the commitments of my own view. To repeat: the suggestion is that, like all desires, the desire-like or motivating belief involved in taking something as one’s reason constitutes an appearance of the good. Now, it is true that I say very little about this in the book. In general, I focus on versions of the guise of the good that appeal to reasons, not to the nature of desire. But the same methodology applies. The proponent of Smith’s line accepts the following principle:

If A wants to φ or is motivated to φ, φ-ing must appear to him in some way good.

2 As argued in Fine (1994). Here, I am indebted to conversations with Cian Dorr.
But there is no brute necessity. So we need to know how this conditional could be explained. In the action theory I defend, motivation by taking as one’s reason is simply non-deviant causation, where non-deviance can be specified without appeal to appearances of the good (Setiya 2007a: 31–2). If this theory is correct, it shows what the examples did not: in Smith’s formulation, ‘that it suffices for intentional action that someone acts on the basis of desire in a [...] stripped down dispositional sense’. The truth of the conditional above would then require an account of appearance on which the conditions for something’s appearing good are inevitably satisfied when one is non-deviantly caused to do it. In the absence of any such account, it is natural to focus, as I did, on the guise of the good for reasons, where there is a tempting line of argument that appeals to Anscombe’s question ‘Why?’ If acting for a reason requires the capacity to give one’s reason, and so to justify one’s action to some degree, we act for reasons under the guise of the good. This familiar picture was my target. In my view, the most plausible version of the guise of the good for desire would follow a similar line, arguing that desires are had for reasons, and that reasons must be seen as good. The latter conception would then fall directly under the argument of the book. At any rate, so I have argued in more recent work (Setiya, forthcoming).

Suppose, then, that we reject the guise of the good for intentional action and the forms of ethical rationalism that depend on it. Suppose, too, that we are sceptical of Velleman’s attempt to extract the standards of practical reason from the idea of practical knowledge. Both Bratman and Smith contend that there are promising forms of rationalism ignored by all of this, and indeed by anything I say in the book. Thus, Bratman argues that a practical commitment theory of intention would support rationalism about some of the norms that govern what we intend: norms of consistency and means-end coherence. And Smith argues that we can extract from the requirement of non-deviance for motivation a more ambitious standard of instrumental rationality.

Before I make a substantive concession, a relatively minor point of response. Smith’s defence of rationalism begins with the problem of deviant causation and the claim that it must be solved by reference to the exercise of a capacity or disposition for instrumental rationality, at least in a local domain. If this is right, at least some standards can be derived from the nature of agency, after all. Although it is very briefly done, however, Reasons without Rationalism does propose an alternative theory of non-deviant causation (Setiya 2007a: 31–2). For basic action, non-deviance consists in one’s behaviour being guided towards the satisfaction of desire in much the way that operations of non-rational creatures can be guided towards an end; since this

process is not mediated by belief, it does not depend on instrumental rationality. For non-basic action, non-deviance consists in doing something in the way one planned. And for intentions and desires, non-deviance consists in forming one attitude on the basis of another because one is so disposed. The latter condition now seems to me too permissive. Inspired by Yablo (1992) on ‘proportionality’, we might regard the transition from C to E as non-deviant if and only if it is the exercise of a disposition to go from C to E, and one does not have this disposition merely in virtue of being disposed to go from some more inclusive condition, C*, of which C is a determinate, to E. In that sense, or perhaps some strengthening of it, the specific content of the mental state is relevant to motivation. Still, the content need not be relevant in the normative sense that the disposition exercised is even approximately rational. As I argue in the book, I can decide to drink coffee because I love Sophocles, or to put money in a pencil sharpener because I want a drink from a nearby vending machine – where this is the ‘because’ of motivation or non-deviant causality – so long as I am appropriately disposed. Such dispositions are unusual, but their presence would ensure that the relevant transition is not, in Smith’s terms, a ‘fluke’.

This reply is, however, superficial. Even if I am right that we can distinguish motivation from causal deviance without appeal to the exercise of (instrumental) rationality, there is room for a kind of rationalism neglected in my book and implicit in Smith’s remarks. Suppose that, in order to have psychological states of a given kind, one must roughly conform to corresponding standards, which govern combinations of and transitions among those states and others. This is part of what it is to have beliefs, desires, intentions and the rest. Since they figure in the nature of these states, it may be argued, the standards in question count as standards of excellence for one’s dispositions of motivation and thought. Alternatively, they are standards of proper psychological functioning, where the sense of ‘function’ is not that of artefacts or living things, but of functionalism in the philosophy of mind. The dispositions one must approximate in order to have an attitude set standards of rationality for its formation and revision. This applies to the attitudes that figure in practical thought, regardless of the nature of non-deviance. One could think of Bratman’s proposal as an instance of such rationalism that appeals to the functional role of intentions-as-plans.

4 This is true even if Smith (2004, §IV) is right to argue that, in order to perform a basic action, one must believe that it is something one can just do, since this belief plays no essential role in the story of guidance.

5 Setiya (2007a: 63–6), and for the bearing of these claims on the prospects for ethical rationalism, Setiya (2007a: 94–8). Note that the conditions indicated in the text are not sufficient for intending or wanting something for a reason, any more than the mere exercise of dispositions is sufficient for acting on the ground that p: one must take that consideration as one’s reason for acting.

6 See also Wedgwood (2008: 1134–5).
It is a striking feature of rationalism, in this guise, that it need not conflict with the virtue theory of practical reason. Its target is the idea of a good psychological disposition, and it does not attempt to find distinctive standards for dispositions of practical thought among our traits of character, or for traits of character in our psychology. The standards of reason and virtue may be the same: standards of proper functioning that derive from the nature of the mind.

That the neglected view may not conflict with the virtue theory partly explains its omission in *Reasons without Rationalism*. But since I do not accept it, I need to say why. Bratman anticipates the principal objection: that the view has unacceptable implications for what there is reason to do. For Bratman and Smith, as for Dreier (1997) and others, the standards most likely to follow from the nature of psychology are ones of efficiency in the satisfaction of desire and coherence in matching means to ends. After all, it does seem part of the functional role of intention and desire to make one act so as to achieve their objects in the light of one’s beliefs. What is taken to be less clear is whether practical rationality is exhausted by its instrumental forms, whether there are further requirements to add. Against this, I would argue that means-end coherence and efficiency are not just limited or partial, as conceptions of good practical thought, but in fact conflict with it. Aspects of this argument appear in two papers separate from the book: ‘Cognitivism about Instrumental Reason’ (Setiya 2007b) and ‘Is Efficiency a Vice?’ (Setiya 2005). In the latter, I argue that means-end efficiency in desire is a defect of character and therefore cannot be required for practical rationality in the sense that connects with reasons. In the former, I give a version of the argument that Bratman recounts in his comments: if means-end coherence for intention were a principle of practical reason, suitably stubborn intentions would provide us with reasons for intending the necessary means, even when their objects are vicious. (In fact, I argue for the more disturbing consequence that these reasons would be decisive; Setiya 2007b: 660–2.)

Each of these arguments depends, more or less explicitly, on the principle I call ‘*Reasons*’, according to which the fact that \( p \) is a reason for \( A \) to \( \phi \) just in case \( A \) has a collection of psychological states, \( C \), free of false beliefs, such that the disposition to be moved to \( \phi \) by \( C \) and the belief that \( -p \) is a good disposition of practical thought. As Bratman points out, my arguments also assume a ‘simple isomorphism between norms of practical rationality on intention, on the one hand, and, on the other hand, norms of good practical thinking and good dispositions of practical thought’. Against this, Bratman insists, ‘we should think of the demand for consistency and coherence of intention as only one element in an account of what good practical thinking

---

7 It counts as a form of rationalism by the definition at Setiya (2007a: 14–15), despite occasional carelessness, as at 85–6 or 106–7, which focus on forms of rationalism that do reject the virtue theory.
is, in the sense of good practical thinking that can do the needed work in a principle like *Reasons*. Smith is sceptical of *Reasons* itself, contrasting it unfavourably with his own ‘advice model’ of the relationship between practical reasons and practical rationality.

So, the situation is this. If *Reasons* is correct and there is a simple isomorphism between norms of rationality and good dispositions of practical thought, Bratman’s rationalism about means-end coherence has unacceptable implications for what there is reason to do. His response is to deny the antecedent, by denying the isomorphism. A similar conditional holds for Smith on instrumental rationality. If *Reasons* is correct, and means-end efficiency is a good disposition of practical thought, there is reason to act on even thoroughly vicious desires. But Smith denies the antecedent, since he denies the truth of *Reasons*. Can anything be said in defence of my assumptions, and thus on behalf of my schematic arguments against Bratman and Smith?

Begin with Bratman. A *prima facie* worry about his line is that, while it is easy to see how an element of practical rationality could be supplemented in a complete account, it is much harder to see how good practical thought could violate genuine norms of practical rationality – which is what Bratman needs in order to block the implications I dislike. One way to develop his picture so as to make sense of this possibility is to suppose that, just as we have *pro tanto* reasons and the all things considered ‘should’, where the latter is a product of the former, so good practical thought, in the sense that figures in *Reasons*, is a product of *pro tanto* norms of practical rationality, somehow combined. There is nothing paradoxical in the idea that acting as I should involves acting against a *pro tanto* reason, one that is outweighed by others. Likewise, we might say, there is nothing paradoxical in the idea that good practical thinking goes against a genuine *pro tanto* norm of practical rationality, when other norms are more insistent.

Now, in the specific case of consistency and means-end coherence for intention, it is not clear that appeal to a *pro tanto* norm captures the peremptoriness of the relevant demands: one *must* intend the necessary means to one’s ends. But this is controversial, and may not generalize.

More pressing is how to make sense of the further layer of balancing in the present account. When we talk about the weight of evidence, the natural picture is that belief comes by degree and degrees of evidence turn on the degrees of belief appropriate in light of one’s evidence. Similarly, when we talk about the weight of practical reasons, the natural picture is that desire and motivation come by degree, and the weight of a reason turns on the strength of desire or motivation appropriate in light of it. In each case, the

---

8 In his comments, Smith mainly objects to *Reasons* left-to-right. The argument in the text does not depend on this: it applies the principle of *Reasons* right-to-left. As I understand it, however, Smith’s advice model conflicts with both conditionals.

9 Setiya (2007b: 653), citing, ironically, Bratman himself!
weight is understood as a function of two things: a state that comes by degree – belief, desire, motivation – and a standard that determines the right degree to have. Are we to suppose a similar structure for pro tanto norms of practical rationality, on which the weights that go into the balance that determines good practical thought correspond to degrees of something – what could it be? – which are correct or appropriate in light of some further, unspecified standard? More abstractly: if norms of rationality derive from the functional role of psychological states like intention, what could explain their relative weights? There is nothing in the structure of the view from which they could derive.

Suppose, then, that we accept the simple isomorphism, and map norms of rationality directly on to good dispositions of practical thought. Still, my arguments against Bratman and Smith would fail if Reasons were false. But how could it be? There must be an intimate relation between practical reasons and practical rationality. And it must be specified so as to avoid the ‘problem of imitation’. We cannot identify reasons with considerations that would move us if we were fully rational, since being fully rational might change our reasons in significant ways. Reasons accommodates this fact (Setiya 2007a: 9–14). If we hope to reject it, we must turn to the advice model, which also solves the problem of imitation, but blocks the argument from means-end coherence as a requirement of rationality to reasons for intending the means to stubborn but vicious ends, and the parallel argument for means-end efficiency in desire.

The main difficulty for the advice model is that, as Smith acknowledges in The Moral Problem, the ‘relative’ version – ‘there is a reason for A to φ in C just in case A would want himself to φ in C if he were fully rational’ – conflicts with the universality of reasons – ‘if there is a reason for A to φ in C, there is a reason for anyone to φ in C’ – unless there is convergence in the relevant conditional desires of fully rational agents (Smith 1994: 164–74).10 Hence, the ‘non-relative’ version, according to which there is a reason for A to φ in C just in case everyone would want themselves to φ in C if they were fully rational. But this is exceptionally demanding: it makes reasons very hard to come by. That is a standard complaint. What has not been sufficiently emphasized is that it leads to something worse. If convergence fails for the desire to φ in C, the non-relative advice model implies that an agent who finds herself in C and is moved to φ on the basis of fully rational deliberation from a correct conception of her circumstance has no reason whatever to do so! That is incredible. We might accept it, I suppose, if there were no alternative, no way to picture the relationship between reasons and

10 Full rationality is said by Smith (1994: 156) to have three components: no false beliefs, all relevant true beliefs and correct deliberation. Only the last of these corresponds to practical rationality or good practical thought as I understand them, but since ignorance and false belief will not be relevant here, we can set this difference aside.
rationality that could avoid this dubious result. But there is: we can map the norms of rationality on to good dispositions of practical thought and appeal to *Reasons*.

Smith’s direct objections to that principle strike me as inconclusive. He gives examples in which a consideration seems to be a reason for A to φ even though it does not satisfy *Reasons*, since being moved by it is in some way self-defeating. Smith’s first example, in which the fact that I believe that I am Jesus is a reason to seek help, might be disputed on the ground that we *can* recognize irrational beliefs and act accordingly. But there are simpler ways to illustrate his point. Suppose, for instance, that I have forgotten a meeting at noon. Won’t that count as a reason to check my calendar, even though it would not make sense to be moved by the corresponding belief? If I knew that I had forgotten my noon appointment, there would be no need to check. My response is that, in situations like this, the cited consideration is *not* a reason for action, but a related reason easily confused with it. That I have forgotten my noon appointment cannot justify checking my calendar. Instead, it is a reason for me to be glad if I do, and dismayed if I do not. Compare discussions of ‘the wrong kind of reason’, as when the fact that it would make me happy to believe in God seems to be a reason for that belief, or in Kavka’s (1983) toxin puzzle, where the benefits of having an intention seem like reasons to have it, even though there will be no reason to act on that intention when the time comes. Here, again, I think we are conflating reasons for one thing with reasons for another (Parfit 2001). That it would benefit me to intend some pointless action, but not to do it, is a reason to wish for that intention, not a reason to intend. That it would make me happy to believe in God is a reason to desire that belief, not a reason to believe. These considerations fail to count as reasons for intention or belief precisely because it is bad reasoning to form the relevant attitude on the basis of them – though perhaps good reasoning to form the corresponding desire. Smith’s second example, of spontaneous action, is controversial in a different way. For if spontaneity is a virtue, and the virtue theory is true, the fact that an act is spontaneous *will* imply that there is a reason to do it: not the fact that it would be spontaneous, but the fact by which the spontaneous person would be moved. We can agree with Smith that there is a reason here without accepting his view of what the reason is, and thus without endangering *Reasons*.

A final question remains. Even if I am right that there are serious obstacles to the form of rationalism proposed, in different ways, by Bratman and Smith, how does this constitute a response to the arguments that motivate it? Why don’t norms of coherence follow from the nature of intention, on a practical commitment view? Or from the functional roles of other psychological states?

The answer, I think, is that if such norms can be extracted, as perhaps they can, they are not best understood as norms of practical rationality in a sense
that relates, even at one remove, to reasons. We need a distinction that figures in Davidson’s thinking about the ‘constitutive ideal of rationality’, from which the present form of rationalism loosely derives.\textsuperscript{11} According to Davidson:

The sort of irrationality that makes for conceptual trouble is not the failure of someone else to believe or feel or do what we deem reasonable, but rather the failure, within a single person, of coherence or consistency in the pattern of beliefs, attitudes, emotions, intentions, and actions. Examples are wishful thinking, acting contrary to one’s own best judgement, self-deception, believing something that one holds to be discredited by the weight of the evidence. (Davidson 1982: 170)\textsuperscript{12}

Davidson notably does not include: acting contrary to the reasons of which one happens to be aware, or believing something that is in fact discredited by one’s evidence. At least in this passage, he seems to distinguish the norms constitutive of psychology from the standards for being reasonable that are the primary objects of ethical and epistemological dispute. In any case, that is the sort of distinction I need. Normativity comes in different kinds, only some of which relate immediately to reasons. Think of the contrast between consistency in belief, both logical and probabilistic, and epistemic justification. Turning from epistemology to ethics, Bratman is right to think that the norms potentially derived from his practical commitment view do not map directly on to good dispositions of practical thought. Where he goes wrong, I believe, is in suggesting that good practical thought is still to be explained in terms of them. Likewise, Smith may be right to find standards of proper functioning in the nature of psychology, but wrong to think that in exploring them, we are doing ethics. The norms implicit in psychology are not those of good practical thought. What I say in my book about action theory also goes for the philosophy of mind: we need to study it in ethics not because it provides the foundation for practical reason, but in order to see that it does not.\textsuperscript{13}

\textit{University of Pittsburgh}

\textit{Pittsburgh, PA 15260, USA}

\texttt{kis23@pitt.edu}

\textsuperscript{11} Compare Setiya (2007a: 65–6, 98), where this distinction was ignored.

\textsuperscript{12} See also Davidson (1985). He seems to forget his own distinction in later work, as when he writes about convergence in Davidson (1995: 49–51).

\textsuperscript{13} Along with Michael Bratman and Michael Smith, I am grateful to Ralph Wedgwood, to participants in Michael Smith’s Seminar in Systematic Ethics at Princeton University and to an audience at the Eastern Division Meeting of the American Philosophical Association in Philadelphia, December 2008.
References


