

This is the penultimate draft of a critical study published in *Philosophical Quarterly* 60, no. 240 (July 2010). Please cite the published version.

THE STANDARDS OF PRACTICAL REASONING

Matthew Silverstein

Reasons without Rationalism. By Kieran Setiya. (Princeton UP, 2007. Pp. ix + 131. Price £23.95.)

1.

Kieran Setiya's target in this insightful, challenging book is a view he labels *ethical rationalism*. Essentially, rationalism is the idea that good practical thought can be understood apart from ethical virtue. It is, in other words, the view that practical reasoning is governed by standards that are *internal to it* and that may, therefore, be distinct from the standards of good character. The rationalist's project then, is to attempt to derive the standards of practical reasoning from an account of the nature of practical thought in general.

Setiya believes this project is hopeless. He rejects the notion that the nature of practical thought is sufficiently robust to generate standards that are internal to practical reasoning. In rationalism's place Setiya offers what he calls a *virtue theory of practical reason*. According to this virtue theory, there are no standards for practical reasoning independent of the standards of good character. Instead, "being good as a disposition of practical thought [just] is being a disposition of practical thought that is good as a trait of character" (8). Take, for example, the disposition to assign overriding weight to considerations of self-interest when deliberating about what to do. Setiya's virtue theory entails that we cannot determine whether this is a *good* disposition of practical thought—and thus whether assigning overriding weight to considerations of self-interest is a feature of good practical reasoning—without considering this disposition's place in a broader account of the virtues of good character. More specifically, this disposition counts as a good disposition of practical thought if and only if it is included in that broader account of good character. Despite the rationalists' insistence to the contrary, then, the standards of practical reason cannot be understood apart from the standards of good character.

One of the few disappointments of Setiya's book is that he never develops his view into what we expect a *virtue theory of practical reason* to be, namely, a detailed account of the character traits that constitute good practical thought. Setiya acknowledges that he has no such account to offer (at least in this book). Instead

his interest lies merely in establishing that the standards of practical reasoning cannot float free from the virtues of character—*whatever those virtues may be*. Yet this is no trifling conclusion. To see this, consider a familiar question in moral philosophy: Do we have reason to be moral? That is, do the standards of practical reasoning require one to give weight to considerations that figure prominently in the thought of people of good character. Rationalists give different answers to this question, but they are united in treating it as a legitimate topic of concern. After all, if practical thought is governed by standards internal to it (as the rationalists suppose), then it is reasonable to wonder how those standards are related to the standards of good character. If, however, rationalism is false, then worries about the place of morality in practical reasoning will have been misplaced. For according to the virtue theory, “if a virtuous person would be moved by certain considerations, it *follows* that they count as reasons to act” (3). The upshot of Setiya’s conclusion, then, is that vast tracts of moral philosophy really do rest on a mistake.

2.

Setiya’s argument for this conclusion involves several controversial steps. One concerns the metaphysics of *goodness*. Setiya observes that things are often evaluated *as members of a certain kind*. When we say, for example, that something is a good knife, we mean that it is good *as a knife* (and not merely that it is a knife and that it is good). One feature of this sort of evaluation is that “the standards for being a good *F* may be different from the standards for being a good *G*, even when *Fs* are a kind of *G*” (80). This is familiar enough: even though a parent is a kind of person, someone might be a good parent without being a good person (84). Setiya’s next moves are more contestable, however. First, he suggests that when we evaluate dispositions of practical thought, we do so in the manner Setiya has identified—that is, *as dispositions of practical thought*. Second, he claims that when standards for evaluation come apart as they do in the case of the bad person who is a good parent, there must be something in the nature of the specific kind in question (in this case, parenthood) that accounts for this evaluative difference. Put more schematically, “if *Fs* are a kind of *G*, and being a *good F* is not simply a matter of being an *F* and being a *good G*, there must be something in the distinctive nature of *Fs* to explain or illuminate their goodness” (83).

These two claims are far from obvious. However, the conclusion they jointly entail—namely, that “dispositions of practical thought must be distinctive, as traits of character, in a way that explains why they are subject to their own evaluative standard”—is one that most rationalists embrace (85). It is precisely because rationalists hold that there *is* something distinctive about practical reasoning that they ascribe specific standards of goodness or correctness to it. Therefore, although I can imagine rationalists who would resist Setiya’s account of the metaphysics of goodness, I shall offer no such resistance here.

3.

Setiya's account of the metaphysics of goodness forms the basis of his argument against rationalism, for he is convinced that "there is *nothing* about our dispositions of practical thought that could ground the difference in evaluative standards" posited by the rationalist (84–5). The argument for this claim occupies the bulk of the book and involves a lengthy detour through the philosophy of action. Setiya suggests that rationalists' attempts to identify something distinctive about practical thought that could ground standards internal to it all presuppose a "normative conception of agency" according to which action for a reason is always performed under the guise of the good—or, more generally, under the guise of the *normative* (16). Thus Setiya devotes the first part of the book to defending a novel theory of intentional action according to which action for a reason is *not* performed under the guise of the normative. Crucially, Setiya's theory entails that "there is *nothing* in the nature of action, or of practical thought, from which the standards of practical reason could derive" (18).

Setiya's attack on normative conceptions of agency begins with an appeal to intuitions. He knows that one way to discredit the thesis that action is performed under the guise of the normative is to provide counterexamples in the form of cases of clear-eyed akrasia—cases where agents intentionally pursue what they recognize to be worthless or even bad. I join Setiya in finding such counterexamples—he introduces many—"entirely convincing" (38). He recognizes, though, that philosophers with contrary intuitions about action are likely to react differently, and so he offers a second, more systematic argument against the guise of the normative thesis.

This argument begins with the observation, first made by Elizabeth Anscombe in her 1957 book *Intention*, that when one acts intentionally, one has a special sort of knowledge—"knowledge without observation"—of one's own behavior. Setiya actually focuses on a simpler phenomenon, which he labels *Belief*: "when someone is acting intentionally, there must be something he is doing intentionally, not merely trying to do, in the belief that he is doing it" (26). Setiya remarks that "the central puzzles of action theory" all revolve around this phenomenon, and hence that giving "a plausible account of this connection is a condition of adequacy on a philosophical theory of acting for reasons" (27). After all, "a necessary truth cannot be mere happenstance" (27).

Setiya then argues that normative conceptions of agency make a mystery of *Belief*. The claim that my acting intentionally requires me to take some consideration as favoring my action does not entail that I believe I am performing that (or any) action. In Setiya's words, the thesis of "the guise of the good does nothing to help explain *why* action that is done for reasons, or intentional action in general, must

satisfy *Belief*^o (41). Thus, if we can develop a theory that *does* explain *Belief*, we will have sufficient reason to reject the normative conception of agency.

In the course of developing such a theory, Setiya acknowledges that intentional action involves one's taking something to be a reason, but he urges that the sense of "reason" at work here is *explanatory*, not *normative*. When one acts intentionally, one must have a belief not about the *justification* of one's action, but rather about *the psychological explanation* of one's behavior. Moreover, this belief must play a role in motivating that behavior. After considering various details and provisos, Setiya arrives at the following formulation: "to take *p* as one's reason for doing ϕ is to have the desire-like belief that one is *hereby* doing ϕ because of the belief that *p*" (46). This is the core of Setiya's alternative theory of action. As he observes, for all its intricacy this view is actually a kind of "minimalism": compared to most other accounts of agency, it places few constraints on rational agents and their actions (98).

Setiya's account of intentional action constitutes an original and important contribution to the philosophy of action. Nevertheless, I have some qualms about the theory and the arguments supporting it.

First, notice that on Setiya's view, my taking something to be my reason for ϕ -ing presupposes that I am *already* ϕ -ing: "When I go for a walk because the weather is fine, I am motivated by a state that is at once the belief *that I am walking outside* for that reason, and like a desire in causing me to do it (and to do it for that reason)" (42, emphasis added). But this means that the sort of taking-to-be-a-reason at work in Setiya's theory of agency cannot be the sort of taking-to-be-a-reason that figures in *practical reasoning*—that is, in the deliberation that *precedes* action.

Why does this matter? After all, Setiya is interested in *acting for a reason*. And he claims that a theory of *acting for a reason* provides a story not about how action comes about, but rather about *what is going on when someone is acting*. If this is correct, then the sort of taking-to-be-a-reason that happens when someone is acting is the only sort that should concern Setiya, and he may therefore deny that my qualm constitutes an objection. However, this would be to forget what is at stake in the larger argument against rationalism. Recall that rationalists believe that the nature of practical thought is sufficiently robust to ground standards internal to it. Setiya attempts to refute this by showing that the sort of taking-to-be-a-reason that figures in practical thought is not the normative (or otherwise robust) sort required for rationalism. Yet now it looks like he and the rationalists are describing two different phenomena. For when the rationalist attempts to extract a theory of reasons for action from the nature of practical reasoning, he focuses on the kind of reasoning that *gives rise to actions*—that is, the kind of reasoning we engage in when we are deciding what to do. And for all that Setiya's

theory tells us, it may still be the case that *this* kind of reasoning is necessarily carried out under the guise of the normative. A theory about the attitudes or judgments that necessarily accompany action does not tell us anything about the attitudes or judgments involved in reasoning about what to do. The rationalist is concerned with the latter; thus it is not clear how Setiya's account of the former is relevant to an argument against rationalism.

A second qualm concerns the explanatory adequacy of Setiya's theory of action. Recall that the normative conception's failure to explain *Belief* is Setiya's primary reason for rejecting that view. Setiya solves this explanatory puzzle by positing that taking something to be a reason is just a matter of having a certain belief about one's own behavior. Yet Setiya also wants to capture the familiar sense in which taking something as a reason moves us to act, and so he concludes that the belief in question is *desire-like*: "It is a belief-like representation of *p* as my reason to act, and at the same time a decision to act on that reason, something by which I am led to do so" (39). But by appealing to the notion of a desire-like belief, Setiya solves one explanatory puzzle by creating another. That is, he solves the puzzle of *Belief* in a way that makes it hard to understand *why* taking something as a reason would give rise to action. After all, beliefs do not normally have that power. Setiya seems to think that by describing the belief as a desire-like one, he is discharging this explanatory burden: if the belief is desire-like, then surely that explains why it motivates action in ways that other beliefs do not. But this is not an explanation; it is merely a redescription of the phenomenon that needs to be explained. Here is Setiya's original characterization of that phenomenon: "taking something as one's reason is both desiderative or motivational *and* cognitive"; it is "a state that has features of both" beliefs and desires (40). The pressing question is *how* one attitude can have both sets of features. Like standard beliefs, the belief in question "represents its content as being true—after all, I know what my reasons are" (40). But then why would that belief have "the power to cause or motivate the action it depicts" when other beliefs lack that power (40)?

Notice that the normative conception of agency has *less* trouble meeting this explanatory burden. According to the normative conception, the belief at work in taking something to be a reason is a *normative* belief. And it is commonly (though hardly universally) accepted that normative beliefs or judgments necessarily motivate. It is one of the features that marks them as normative. Thus, the normative theory of agency can explain the desire-like features of these beliefs by appealing to their *content*. This particular explanatory avenue is not open to Setiya, however. On his view, the content of the belief that constitutes taking something as a reason is purely psychological; it is a belief merely about the psychological explanation for one's action. And it is unclear why some such beliefs would have the power to motivate, especially since other such beliefs evidently do not.

This leads me to conclude that the explanatory advantages of Setiya's theory of action over the normative conception of agency are far from clear. The best theory of action for a reason would account for *Belief* without leaving any mystery about how the attitudes involved in taking something to be a reason could motivate action. And there is a prominent view in the literature that purports to do just that: J. David Velleman's constitutive aim approach. Like Setiya, Velleman thinks that when one takes something to be a reason, one has a belief about the psychological explanation of one's action, and so Velleman too has a ready solution to the *Belief* puzzle. His theory includes an additional ingredient, though—one that explains how these beliefs give rise to action. Velleman contends that it is constitutive of agency that agents have a *motive* or *drive* to understand their own behavior in certain ways. Beliefs about these sorts of explanation engage this motive, and that is why they (and only they) play a special role in bringing about action.

Velleman's theory is no friendlier to the thesis that action for a reason is performed under the guise of the good than Setiya's, and so one might think that Setiya could salvage his argument against rationalism by retreating to something like Velleman's view. But as Setiya recognizes, once we add a motive or drive to our theory of agency, we reach an account according to which action has a constitutive aim. And a constitutive aim of action looks like precisely the sort of thing from which we should be able to derive an internal standard of correctness for action. Put another way, once we introduce the notion of a constitutive aim, we arrive at a plausible basis for ethical rationalism.

4.

That we can arrive at a rationalist view without embracing a normative conception of agency is somewhat surprising, at least in light of the way Setiya outlines the argument of his book. He frequently suggests that what dooms ethical rationalism is simply its dependence on the thesis that intentional action is performed under the guise of the good.

To be fair, much of the second half of Setiya's book is devoted to showing that various versions of rationalism do presuppose a normative conception of agency. This turns out to be fairly easy with respect to certain rationalist views, such as recognitional theories according to which the subject matter of practical thought is just thinking about good reasons as such—that is, about what is best to do, or about how to live the good life. Matters prove to be only slightly more complicated when Setiya turns to the Kantian variant of ethical rationalism defended by Christine Korsgaard in *The Sources of Normativity*. Korsgaard hopes to derive standards of practical thought from a theory of agency. Setiya argues persuasively, though, that Korsgaard's conception of agency is at heart a normative one. Central to her account of agency is the notion of a *practical identity*, which she

characterizes as “a description under which you find your life to be worth living and your actions to be worth undertaking” (101). On Korsgaard’s view, one cannot act for a reason without embracing a way of life as *worthwhile*. So, at least two prominent varieties of rationalism do rely on normative conceptions of agency and are therefore vulnerable to Setiya’s arguments against the idea that actions are performed under the guise of the normative.

As we have seen, though, there are varieties of ethical rationalism that do *not* presuppose a normative conception of agency. These views attempt to ground standards of practical thought in accounts of agency that reject the guise of the normative but that are nevertheless more robust than Setiya’s minimalist theory of action. Velleman’s constitutive aim approach is one such view. Another is the Humean view according to which we can derive instrumentalist standards of practical reasoning from the desire-based nature of rational motivation. Setiya’s arguments against the guise of the normative provide no leverage against these additional forms of rationalism, so in the final sections of his book he offers more specialized arguments directed at the accounts of agency underlying these two approaches.

Setiya’s argument against instrumentalist versions of rationalism revolves around his rejection of the Humean theory of motivation upon which such views rest. This discussion is interesting and important, but I shall not address it here, since I do not find an instrumentalist account of practical thought plausible. I shall focus instead on Setiya’s criticism of Velleman’s approach.

Setiya’s and Velleman’s theories of agency have much in common. Both eschew the thought that action is performed under the guise of the normative, and both maintain instead that the reasons for which we act are *explanatory*, not *justificatory*. Moreover, both Setiya and Velleman believe that action is primarily about knowing what one is doing, and therefore that action aims at self-knowledge. On Velleman’s view, however, action aims at self-knowledge in a more robust sense than it does on Setiya’s view. For Velleman, the aim is embodied by a drive to understand one’s own behavior—a drive that can be more or less satisfied by different instances of behavior. As I noted above, it is this feature of Velleman’s account of agency that provides him with the materials for a rationalist theory of the standards of practical reasoning.

The evident similarities between their respective approaches naturally give rise to the following question. Given that Setiya is willing to move so far in the direction of Velleman’s view, why does he balk at going all the way? What in Velleman’s account of action—aside from its compatibility with ethical rationalism—gives Setiya pause?

Setiya's primary objection is that Velleman's expanded theoretical apparatus is superfluous. Recall that for Setiya, the central puzzles of action theory all revolve around the phenomenon of our knowledge of our own behavior. According to Setiya, though, his own theory of agency shows that we do not need to posit a desire for self-knowledge in order to explain this phenomenon, and thus "there is no need to appeal to this [desire] in the explanation of intentional action" (109). In other words, given that Setiya's minimalist theory provides all the explanation we need, there is no motivation for Velleman's more robust account.

This claim strikes me as mistaken in at least two ways. First, as I argued above, Setiya overestimates the explanatory power of his own view, especially in comparison with Velleman's. Even if Setiya is right that "we can make good sense of desire-like belief, in its own right"—even if we have no reason to rule out the possibility of such an attitude—it hardly follows that there are no mysteries left to explain. As I suggested above, we might reasonably wonder why some beliefs are desire-like while others are not. Setiya responds by insisting that the desire-like beliefs are actually "a distinctive sort of psychological state" (109): they are *intentions*, and *that* is why they are desire-like. But this response does not illuminate the mystery; it merely names it. Velleman's suggestion that action's constitutive aim is embodied by a motive therefore seems anything but unmotivated.

Second, Setiya's assessment of the explanatory power of Velleman's theory rests on the dubious contention that there is only one central puzzle in action theory, namely, the Anscombean phenomenon of spontaneous self-knowledge. This is certainly a central puzzle, but it is hardly the only one. Here are some others. How can events be brought about by agents rather than by other events? Why, from the deliberative point of view, does it feel like we have an open future, even when we accept causal determinism? How can multiple agents share the same intention? If we are going to compare the explanatory virtues of Setiya's and Velleman's theories, we should do so with respect to all of these questions. Velleman has argued that his approach can shed light on the entire lot. And so if Setiya wants to demonstrate that Velleman's expanded theoretical apparatus is unnecessary, he must address more than the phenomenon of self-knowledge. It remains unclear, then, why—once we have accepted their shared idea that the reasons for which we act are *explanatory* rather than *justificatory*—we should refrain from moving from Setiya's minimalist account of action to Velleman's more robust theory.

5.

Where does this leave us? Not quite where Setiya wants us.

One lesson Setiya hopes we will take away from *Reasons and Rationalism* concerns the centrality of action theory to debates about practical reasoning. Here Setiya is

right on the mark. His discussion of rationalist views that presuppose a normative conception of agency demonstrates that most philosophers interested in practical reason have paid far too little attention to action theory—and in particular to the question of whether the theories of agency upon which their accounts of practical reasoning rest pass muster from the point of view of the philosophy of action. That said, I remain unconvinced by Setiya's arguments for his contention that there are *no* accounts of agency that can withstand the scrutiny of philosophers of action and that are *also* sufficiently robust to ground an account of the standards of practical reasoning. Consequently, Setiya's book has had the perverse effect of making me *more* sympathetic to rationalism—or at least to the varieties of rationalism that take action theory seriously, even as it has made me more suspicious of the varieties that do not.

This should not deter potential readers. *Reasons without Rationalism* helpfully reframes the debate concerning the relation between morality and practical reason. Setiya's division of views into rationalist and virtue-theory camps enables us to see the affinities even among seemingly opposing rationalist views, and it forcefully reminds us of the centrality of action theory to discussions of practical reason—and thus to ethics in general. Although I am not persuaded by all of the Setiya's arguments, I have benefited greatly from reading and thinking about them. I am confident that anyone interested in action theory and practical reasoning will as well.

NYU Abu Dhabi