

## Cognitivism about Instrumental Reason\*

*Kieran Setiya*

Whoever wills the end also wills (insofar as reason has decisive influence on his actions) the indispensably necessary means to it that are within his power.<sup>1</sup>

Despite Kant's insistence that the hypothetical imperative "requires no special discussion" since "as regards the volition, [it is] analytic," instrumental reason has come to seem problematic.<sup>2</sup> On the one hand, it seems right to say that one should take the necessary means to one's ends. On the other hand, there are circumstances in which one should not take the means, and even ones in which there is no reason to do so at all.

Imagine that I embark upon on a thoroughly irrational project: I intend to count the blades of grass in my garden.<sup>3</sup> Each day, I wake at six and kneel on the lawn with my magnifying glass, counting away. Despite my intention, however, I do not take what I know to be the

\* In writing this article, I have been much indebted to the inspiration of Michael Bratman's "Intention, Belief, Practical, Theoretical" (forthcoming in *Spheres of Reason*, ed. Jens Timmerman, John Skorupski and Simon Robertson) and to conversations with him at the University of Maryland Workshop on Practical Rationality in April 2005, and at a meeting of the Workshop on Social Ethics and Normative Theory at Stanford University. A reader for *Ethics* sent extremely helpful comments on an earlier draft, and the final version was improved by the advice of the associate editors. Thanks also to Selim Berker, Justin D'Arms, Cian Dorr, Jyl Gentzler, Nadeem Hussain, Agnieszka Jaworska, Niko Kolodny, Jen Morton, Sarah Paul, Nishi Shah, Dan Velleman, Jay Wallace, Ralph Wedgwood, and to audiences at Amherst College, Bowling Green, MIT, and the University of Pittsburgh.

1. Immanuel Kant, *Groundwork for the Metaphysics of Morals*, trans. Mary Gregor (Cambridge: Cambridge University Press, 1998), Ak. 417.

2. *Ibid.*

3. See John Rawls, *A Theory of Justice* (Cambridge, MA: Harvard University Press, 1971), 432.

*Ethics* 117 (July 2007): 649–673

© 2007 by The University of Chicago. All rights reserved. 0014-1704/2007/11704-0003\$10.00

necessary means. Even though I see that I have no chance to complete the enumeration unless I keep track of how many blades of grass I counted on a given day, and where I counted them, I can't be bothered with bookkeeping. So, every morning, I am forced to start again, or at random, or to guess how far I got; and I never manage to complete the count.

If this is how my life is going, something is amiss. The problem of instrumental reason appears in the paradoxical way in which we are prone to describe it. We want to say that I should not devote my life to counting grass; it is a worthless ambition. And so there is no reason at all to keep track of how many blades I counted and where they were. But we also want to say that, if I cannot be persuaded to change my ways, if I am going to count grass come what may, then I should take the necessary means. The puzzle I want to solve in this article is: How can we say these things without contradicting ourselves? How can it be true, at once, that I should take the necessary means to my end—and that I should not?

The example I have used to motivate this puzzle is controversial, as is the content of the “instrumental principle” (Kant’s hypothetical imperative). I address these matters briefly in Section I, arguing that the problem is generated by cases of ordinary clear-eyed *akrasia*. In Section II, I explore the prospects for an account of instrumental reason on which the sense of “should” in which one should take the necessary means to one’s ends belongs to practical reason. In a series of important papers, John Broome has argued that the only sense of “should” at work here is the one that we use in saying what there is most reason, or decisive reason, to do and that the apparent contradiction in the example is removed when we make appropriate distinctions of scope.<sup>4</sup> I argue that this is a mistake. Nor does it help to appeal to a distinction between the “objective ‘should’” of most or decisive reason and the “subjective ‘should’” of practical rationality or good practical thought.

In Section III, I draw, and defend, the surprising implication of this argument, that the instrumental principle is not a principle of practical reason at all. As I go on to explain, the “should” of instrumental reason flows from epistemic requirements on the beliefs that figure in our

4. John Broome, “Normative Requirements,” in *Normativity*, ed. J. Dancy (Oxford: Blackwell, 2000), 78–99, “Normative Practical Reasoning,” *Proceedings of the Aristotelian Society, Supplementary Volume 75* (2001): 175–93, “Practical Reasoning,” in *Reason and Nature*, ed. J. L. Bermúdez and Alan Millar (Oxford: Oxford University Press, 2002), 85–111, and “Reasons,” in *Reason and Value: Essays on the Moral Philosophy of Joseph Raz*, ed. R. J. Wallace, Philip Pettit, Samuel Scheffler, and Michael Smith (Oxford: Oxford University Press, 2004), 28–55; see also Jonathan Dancy, *Practical Reality* (Oxford: Oxford University Press, 2000). Broome has since revised his view, in “Does Rationality Give Us Reasons?” (*Philosophical Issues* 15 [2005]: 321–37); I comment on this briefly in the notes to come.

intentions. It is the “should” of theoretical reason. This is what I mean by “cognitivism” in the title of this article: not that the instrumental principle is itself an object of belief, but an account of instrumental reason as a matter of constraints on the cognitive elements of practical thought; an account of one fragment of what we might think of as practical reason, in theoretical terms.<sup>5</sup>

It is worth noting, finally, that the critical portion of the article, in Section II, and the constructive proposal of Section III, are logically independent. Although the argument against Broome helps to motivate the cognitivist account, one could accept either of them on its own. Thus it is possible to hold, with earlier cognitivists, that the instrumental principle is both practical and theoretical. And it is possible to deny its rational standing altogether. In what follows, I argue against the first response, and I assume that we should hope to avoid the second.

## I

Since the language of “instrumental reason” is ambiguous, we need to begin with some distinctions.

First, there is a contrast between Kant’s hypothetical imperative, which is our topic here, and the “neo-Humean” conception of practical reason as “means-end efficiency”: the disposition to act so as to satisfy one’s final desires. Kant’s principle—the principle of instrumental reason involved in the puzzle I propose to solve—is concerned with taking the necessary means to ends that one intends to bring about. It does not apply to less-than-necessary means or to desires on which one does not (yet) intend to act.<sup>6</sup> It is therefore silent about the balancing of desires and probabilities in practical reasoning and about the best way

5. The idea of a cognitivist account of instrumental reason derives from Gilbert Harman, “Practical Reasoning,” reprinted in *The Philosophy of Action*, ed. Alfred Mele (Oxford: Oxford University Press, 1997), 149–77. It has recently been pursued by R. J. Wallace (“Normativity, Commitment and Instrumental Reason,” *Philosophers’ Imprint* 1 [2001]: 1–26), whose views I address in Sec. III. The use of “cognitivism” in this context is due to Michael Bratman. See “Cognitivism about Practical Reason,” reprinted in his *Faces of Intention* (Cambridge: Cambridge University Press, 1999), 250–64, and “Intention, Belief, Practical, Theoretical,” in *Spheres of Reason*, ed. Jens Timmerman, John Skorupski, and Simon Robertson, forthcoming.

6. As Candace Vogler points out in *Reasonably Vicious* (Cambridge, MA: Harvard University Press, 2002), 154, the restriction to necessary means is quite severe. To say that doing *M* is a necessary means to doing *E*, for a particular agent, *A*, is not to say that it is physically (let alone metaphysically) impossible for *A* to do *E* without doing *M*. But it does imply that everything she could do that is a means to doing *E* involves doing *M*; of the options available to her, doing *M* is part of all those that are ways of doing *E*. This will most often be true when time and resources are limited. I gesture toward the probabilistic generalization of the instrumental principle, briefly, in Secs. III and IV.

to achieve a plurality of potentially conflicting ends. Its application is narrow; it does not purport to be the whole of practical reason.<sup>7</sup>

The second distinction is between the instrumental principle and a principle of transmission for the all-in practical “should,” the “should” of most or decisive reason.

*Transmission:* If you should do *E*, all things considered, and doing *M* is a necessary means to doing *E*, you should do *M*, all things considered, too.

The instrumental principle begins with the fact that one has an end, regardless of whether one ought to have it. It is the claim that one should take the necessary means to whatever ends one intends to bring about. The principle of transmission, by contrast, begins with an action that one should perform. It derives a “should” from another “should.”

The proper formulation of the instrumental principle is a matter of dispute. Different theories of instrumental reason support different principles, qualified so as to appeal to an agent’s beliefs about the necessary means or to the agent’s belief that intending to take the means is itself a necessary means. We will come back to this as we go on. But we can state the basic problem of instrumental reason without worrying about the details. The difficulty raised by the grass-counting example is that, when we say that I should take the necessary means to my end—I should keep track of what I’ve done so far, if am going to count the grass—we cannot be using the all-in practical “should,” at least not in the obvious way, since there is no reason for me to take the means. There is no reason to keep a daily record of my work, and I should simply give it up.<sup>8</sup>

7. By contrast, the trait of efficiency—in the sense of being disposed to satisfy the balance of one’s desires in the light of one’s degrees of belief about the means to doing so—might be proposed as a general conception of practical reason. If we allow for “constitutive” as well as “productive” means, and adopt the broadest possible conception of desire, as anything that belongs to an agent’s “subjective motivational set,” we then come close to Bernard Williams’s theory of “internal reasons,” on which reasons to act can always be traced to an agent’s prior desires. (See “Internal and External Reasons,” reprinted in his *Moral Luck* [Cambridge: Cambridge University Press, 1981], 101–13.) This is one kind of “instrumentalism,” but it is not the instrumental principle to be discussed below. On the distinction between the hypothetical imperative and the idea of balancing among desires in general, see Christine Korsgaard, “The Normativity of Instrumental Reason,” in *Ethics and Practical Reason*, ed. Garrett Cullity and Berys Gaut (Oxford: Oxford University Press, 1997), 215–54, and “The Myth of Egoism” (Lindley Lecture, Department of Philosophy, University of Kansas, Lawrence, 1999).

8. For examples of this kind, see Michael Bratman, *Intention, Plans and Practical Reason* (Cambridge, MA: Harvard University Press, 1987), 23–27; Broome, “Normative Requirements,” 89–90, and “Reasons,” 29–30; Dancy, *Practical Reality*, 42–43; Wallace, “Normativity, Commitment and Instrumental Reason,” 15–16; and Joseph Raz, “The Myth of Instrumental Rationality,” *Journal of Ethics and Social Philosophy* 1 (2005): 2–28. (As I understand

The description of this case is not as strange as it may sound. One can accept it without saying that there is never a reason to take the necessary means unless there is reason to pursue the end, so long as there are some cases in which there is no reason to take the means. Nor does it conflict with instrumentalism in the “neo-Humean” sense. If all our reasons stem from our final desires, and if the intention to count blades of grass answers to no final desire (and does not itself amount to a final desire), then, for the instrumentalist, there is no reason to act on it and, thus, no reason to take the necessary means. Finally, one can say that there is no good reason to take the means, nothing that would tend to justify doing so, without denying that I might, say, decide to buy a notebook, precisely on the ground that I am counting grass and need one in order to do so accurately. My intention for the end can supply my reason for taking the means without there being any consideration that counts in favor of doing so: no reason in the normative sense.

Nevertheless, it is worth showing that the problem of instrumental reason can be generated on more modest grounds. Suppose, then, that intending an end always provides some reason for one to take the necessary means.<sup>9</sup> When I intend to count the blades of grass in my garden, it is a mistake to say that there is no reason for me to keep track: a reason is given by the fact that I am engaged in grass counting and that I need to keep track of the count in order to complete it. Still, this can't explain the sense in which I *should* keep track of the count. From the fact that there is a reason to do something it does not follow that I should do it, since the reason may be outweighed, as it surely is in the case at hand. Appeal to my intention for the end as generating a defeasible or *pro tanto* reason is thus too weak: it fails to capture the requirement expressed by the instrumental principle. It is not just that there is reason to take the necessary means to one's ends; one must do so. The principle of instrumental reason makes a strict or peremptory demand. This is missing from the picture of intentions as providing reasons.<sup>10</sup> (Nor would it help to propose that intentions provide *decisive* reasons. That is simply implausible. It implies that one is always doing what one should do, in acting as one intends.)

It follows that the problem of instrumental reason is raised by any

---

him, Raz is not troubled by such examples because he denies that there is any such thing as the instrumental principle, however it is refined: there is no sense in which we should always take the necessary means to our ends. I won't discuss this skeptical view.)

9. This seems to be assumed by Korsgaard, in “The Normativity of Instrumental Reason,” 215. Although I am conceding it in this paragraph, I think the assumption is false.

10. Similar arguments are made by Bratman (*Intention, Plans and Practical Reason*, 24) and Broome (“Normative Requirements,” 89).

case of clear-eyed *akrasia*. Smoking is pleasant, and so I have a reason to do it—though on the whole, I know that I had better not. Suppose, then, that I decide to smoke, knowing that I need to buy cigarettes in order to do so. Perhaps it is true that my intention adds a further reason, along with the pleasure of smoking, for me to buy them. But these reasons are not conclusive: in the sense of “should” which reports what there is most or decisive reason to do, I know that I should not buy a pack of cigarettes. Doing so would be *akratic*. Nevertheless, if I intend to smoke, it seems that in some sense I should take the necessary means. The instrumental principle applies to me. It follows that the principle cannot be this:

If you intend to do *E*, and you know that doing *M* is a necessary means to doing *E*, you should do *M*, all things considered.

For it is not the case that I should buy cigarettes, in the all-in practical sense of “should.”

The task for a theory of instrumental reason is thus to explain the content of the instrumental principle: to specify the sense in which I should take the necessary means to my ends, even when it is false to say that there is most or decisive reason to do so. What is the “should” of instrumental reason?

## II

A tempting answer is this: the “should” of instrumental reason is the all-in practical “should,” applied in a distinctive way. For Broome, the relevant distinction is one of scope. When we say that you should take the necessary means to your end, our “should” governs a conditional, not its consequent. Thus, according to Broome:

You should [if you intend to do *E* and believe that doing *M* is a necessary means to doing *E*, intend to do *M*].<sup>11</sup>

In other words, you should make the means-end conditional true by conforming to it in one way or another. It does not follow from this wide-scope “should” that you should intend to do *M*, even if, as a matter of fact, you do intend to do *E* and believe that doing *M* is a necessary means to this. Perhaps, instead, you should not intend to do *E*, or it is a matter of indifference whether you respond by taking the means, or giving up the end, or even by revising your belief.

A similar proposal has been made on Kant’s behalf. Thus, Thomas Hill insists that the hypothetical imperative does not conflict with the categorical imperative, even when the necessary means to one’s end are strictly forbidden, “for there is [an] alternative [to taking the means.]”

11. Broome, “Reasons,” 29. I use brackets to resolve ambiguities of scope.

He can abandon the end. Insofar as this remains a possibility, what the Hypothetical Imperative prescribes, in effect, is ‘Take the necessary means or else give up the end.’<sup>12</sup> This is exactly the structure of Broome’s approach.

In developing it, Broome offers a more general account of the practical “should” and of its relation to reasons. He takes the all-in “should” as primitive, reporting what one ought to do, all things considered. No ought fact is inexplicable, however: there is always something in virtue of which one should do whatever it is that one should do. When the explanation takes the form of a single sufficient condition, this condition counts as a “perfect reason.”<sup>13</sup> When an “ought” is explained by a complex of considerations that form a “weighing” structure, some tending to support one action, some another, these considerations count as defeasible or *pro tanto* reasons. Broome attempts to work this idea into a definition of “*pro tanto* reason” in terms of the all-in “should.”<sup>14</sup> The details are not important for us.

What is important is that Broome appeals to a single “should,” the one that we use in describing the balance of reasons (when a “weighing explanation” is appropriate) or the product of a perfect reason (when it is not).<sup>15</sup> It is in this sense that you should [if you intend to do *E* and believe that doing *M* is a necessary means to doing *E*, intend to do *M*]. For Broome, it is mere carelessness to say, as we do when I intend to smoke, that I should buy cigarettes. In the only relevant sense of “should,” I should not do so. But in the very same sense of “should,” I should make true the conditional: [if I intend to smoke and believe that buying cigarettes is a necessary means to this, then I intend to buy cigarettes]. There is no paradox here, even though there is only the all-in practical “should,” since it does not attach to the consequent of the conditional, something I should not do. And it does not in general follow from the claim that I should make true [if *p*, then *q*] and the claim that *p* is true that I should act so as to make true *q*. This form of “detaching” is logically invalid.

Although it can thereby accommodate the problematic cases of grass counting and *akrasia*, I doubt that Broome’s proposal is correct. Nor can it be saved by moving from the “objective ‘should’” of most or

12. Thomas Hill, “The Hypothetical Imperative,” reprinted in his *Dignity and Practical Reason in Kant’s Moral Theory* (Ithaca, NY: Cornell University Press, 1992), 17–37, 24. But see the derivation earlier in Hill’s paper (“The Hypothetical Imperative,” 18), which seems to allow for “detaching” in a way that wide-scope “shoulds” do not. I take up “detachment” in the main text, below.

13. Broome, “Reasons,” 34–35.

14. *Ibid.*, 36–41.

15. In recent work, Broome has backed away from this assumption; see “Does Rationality Give Us Reasons?”

decisive reason to the “subjective ‘should’” of practical rationality or good practical thought. I will explain and defend these claims in an indirect way, by first addressing Broome’s account of what he calls “normative practical reasoning.” The argument will turn on a contrast between this sort of reasoning and the reasoning that corresponds to the instrumental principle.

We begin with a threat of illicit “bootstrapping”: it seems right to say that, in some sense, one should act as one thinks one should. But it would be quite wrong to conclude that, whenever I think I should  $\phi$ , all things considered, I should in fact  $\phi$ . My belief could be false. Broome’s solution to this puzzle is to insist, once again, that the relevant “should” takes wide scope, ranging over a conditional, not just its consequent. His claim is that you should [if you believe that you should  $\phi$ , intend to  $\phi$ ], where this is the “should” of most or decisive reason.<sup>16</sup> Since “detaching” is invalid, it does not follow from this principle, and the fact that I believe that I should  $\phi$ , that I should intend to  $\phi$ , all things considered.

Nevertheless, the problem of bootstrapping has not been solved, as we can see when we recall the means-end transmission of the all-in practical “should”:

*Transmission:* If you should do  $E$ , all things considered, and doing  $M$  is a necessary means to doing  $E$ , you should do  $M$ , all things considered, too.

Consider the following case. I believe that I should  $\phi$ . And there is nothing I can do to change this belief: there are no means available to me for causing the belief to disappear; it is not under my control. It follows that the only way in which I can conform to the conditional [if I believe that I should  $\phi$ , then I intend to  $\phi$ ] is by intending to  $\phi$ . In other words, intending to  $\phi$  is a necessary means to the truth of that conditional. According to Broome, I should make that conditional true, all things considered. It follows, by *Transmission*, that I should intend to  $\phi$ , all things considered.<sup>17</sup> But this need not be so: I may be wrong to believe that I should  $\phi$ , and it may not be true that I should make the corresponding decision. Bootstrapping is not vindicated by the fact that I cannot affect the relevant belief. It remains illicit, in a way that Broome’s principle cannot explain.

This objection may seem impossibly quick; but I think it is sound.

16. Broome, “Normative Practical Reasoning,” 181–82.

17. More generally, means-end transmission supports Patricia Greenspan’s principle that if you should make true a conditional and the antecedent obtains “inalterably,” you should make true the consequent; see Greenspan, “Conditional Oughts and Hypothetical Imperatives,” *Journal of Philosophy* 72 (1975): 259–76, 265.



It hinges on the possibility of a case in which intending to  $\phi$  is a necessary means to the truth of the conditional [if I believe that I should  $\phi$ , then I intend to  $\phi$ ], but in which it is not true that I should intend to  $\phi$ , all things considered. The former claim amounts to this: that there is something I could do that is a means to the truth of the conditional, and everything I could do that is a means to this involves intending to  $\phi$ . (This follows from the general definition of a necessary means: doing  $M$  is a necessary means to doing  $E$ , for a particular agent,  $A$ , just in case there is something  $A$  could do that is a means to doing  $E$ , and everything she could do that is a means to doing  $E$  involves doing  $M$ .) The interpretation of “something I could do” that figures in this account of necessary means may be quite modest: what an agent *can* do, in this sense, is whatever she would do if she made some appropriate decision.<sup>18</sup> So long as there is no decision that would affect my belief that I should  $\phi$ , there is nothing I can do to change that belief, and intending to  $\phi$  will be a necessary means to the truth of Broome’s conditional. What is to prevent this from being the case?

It is no use replying that there must be some way to alter my belief, if only through hypnosis or by tampering with my brain. The validity of normative practical reasoning does not depend on the necessary presence of such means. When they are unavailable, no decision I could make would effectively deploy them, and the objection to Broome will stand. It is in any case implausible to hold that one conforms to the demands of normative practical reasoning when one manipulates oneself in these anomalous ways. They represent a failure, not a success. It is true that Broome’s account does not constrain the means by which one makes his conditional true. But this is not a point in its defense. It is a further reason to doubt that those demands are well expressed by a simple wide-scope “should.”

The proper response to these difficulties is to appeal to what is sometimes called the “subjective ‘should’”: the “should” of practical rationality or good practical thought. The sense in which you should be moved by the belief that there is a reason to  $\phi$  (even when that belief

18. More stringent interpretations of what an agent “can do” would only help my argument, here and later on, by expanding the range of cases in which something counts as a necessary means. For instance, someone might hold that it is not sufficient for being able to  $\phi$  that one would  $\phi$  if one made some appropriate decision; one must also be able to make that decision. I am not sure how to make sense of “freedom of the will” in the sense invoked here, and so I opt for the more modest account in the text. (For related doubts, see Rogers Albritton, “Freedom of Will and Freedom of Action,” *Proceedings and Addresses of the American Philosophical Association* 59 [1985]: 239–51.)

is false), is that being moved in this way is rational or reasonable.<sup>19</sup> Broome wants to assimilate this view to his own account, insisting that the so-called “subjective ‘should’” is just the “objective ‘should’” of most or decisive reason applied with wide scope.<sup>20</sup> But this is not the case. In the subjective sense of “should,” one should be moved to some degree by the belief that there is a reason to  $\phi$ . This cannot mean that one should [if one believes that there is a reason to  $\phi$ , be moved to  $\phi$  to some degree], all things considered, since that would generate the same illicit bootstrapping as Broome’s proposal about the belief that one should  $\phi$ , all things considered. When I cannot affect my belief, the claim that I should [if I believe that there is a reason to  $\phi$ , be moved to  $\phi$  to some degree] entails that I should be moved to  $\phi$  to some degree, by way of means-end transmission. But, as in the previous case, this need not be so. It follows that the “should” of good practical thought is not the “should” of most or decisive reason, applied to conditional claims.

This leaves us with two concepts: *good practical thought* and *reason for action*, which are closely related to one another. Broome is wrong about the character of this relation, but it must be there. I have argued elsewhere that the connection can be stated as follows:

*Reasons:* The fact that  $p$  is a reason for A to  $\phi$  just in case A has a collection of psychological states, C, such that the disposition to be moved to  $\phi$  by C-and-the-belief-that- $p$  is a good disposition of practical thought, and C contains no false beliefs.<sup>21</sup>

In effect, a reason is a premise for an episode of good practical thought whose other conditions are already in place. If the fact that  $p$  is a reason for you to  $\phi$ , then it is good practical thought to be moved to  $\phi$  by a certain array of psychological states, and you have that array—except (perhaps) for the belief that  $p$ . This belief would supply the final material for a good disposition of practical thought. (Trivially, a disposition of

19. I have shifted from the belief that one should  $\phi$ , to the belief that there is some reason to  $\phi$ , because I doubt that it is always part of good practical thought to act on the former belief (as Broome suggests). For discussion of this point, see Alison MacIntyre, “Is Akritic Action Always Irrational?” in *Identity, Character and Morality*, ed. Owen Flanagan and A. O. Rorty (Cambridge, MA: MIT Press, 1990), 379–400; and Nomy Arpaly, “On Acting Rationally against One’s Better Judgement,” *Ethics* 110 (2000): 488–513.

20. Broome, “Normative Requirements,” 94–95. In recent work, Broome rejects the claim discussed in the text, even to the point of questioning the normativity of “rational requirements”—which presumably correspond to practical rationality or good practical thought. (See Broome, “Does Rationality Give Us Reasons?”) According to the argument in the text, Broome’s doubts are correct, at least to this extent: there can be a “subjective ‘should’” without an “objective” correlate.

21. For more extensive discussion, see my *Reasons without Rationalism* (Princeton, NJ: Princeton University Press, 2007), 9–14.

practical thought is good, as such, just in case it is a disposition to engage in good practical thought.)<sup>22</sup>

The last clause of the formula—“no false beliefs”—is required to make sense of one of the cases that first motivates a distinction between the “should” of good practical thought and the “should” of most or decisive reason. In a well-known example (which I have slightly modified), Bernard Williams imagines a thirsty person, presented with what seems to be a glass of cool, refreshing water.<sup>23</sup> In fact, the glass contains odorless petrol. If I am in this situation, is the fact that I am thirsty a reason for me to drink the contents of the glass? As Williams says, the answer would seem to be “no.” If the glass contains petrol, the fact that I am thirsty is no reason to drink from it, at all; there is no good reason to drink what is in the glass. The inclination to say otherwise, to say that I should drink the contents of the glass, turns on the fact that I have a collection of psychological states—including the belief that the glass contains water—such that the disposition to be moved to drink by them, together with the belief that I am thirsty, is a good disposition of practical thought. What the example shows is that good practical thought corresponds to reasons only when it does not rely on false beliefs.<sup>24</sup>

That is why no problem of bootstrapping is generated by the fact that it is a good disposition of practical thought to be moved to  $\phi$  to some degree by the belief that there is a reason to  $\phi$ . If this belief is false, its content cannot be a reason to  $\phi$ , since reasons must be facts. Nor can its role in good practical thought make any other fact into a reason to  $\phi$ , since the practical thought in question would depend on a false belief, and therefore would not correspond to a practical reason. False beliefs about reasons do not illicitly generate good reasons out of nothing.

What we have seen so far is that there is a defect in Broome’s account of normative practical reasoning—it allows for bootstrapping, albeit, perhaps, in fewer cases than a narrow scope view—but that there is a satisfactory replacement. The sense in which one should be moved

22. It is perhaps worth stressing that *Reasons* carries no connotation of “priority” for good practical thought; it is basically symmetric.

23. Williams, “Internal and External Reasons,” 102.

24. A natural question to ask at this point is: why not correct for ignorance as well as false belief? But that “correction” would be a mistake. Reasons for action may correspond to practical thought that depends on ignorance of fact. So, for instance, there may be a reason for the gambler to bet on the horse with the best odds, even though it will lose the race—since he does not know that the horse will lose. His reason corresponds to good practical thought that depends essentially on his being in the dark about that. (For this example, attributed to Frank Jackson, see Dancy, *Practical Reality*, 65–66.)

by one's beliefs about what there is reason to do is that being moved in this way is a good disposition of practical thought.<sup>25</sup>

Things look different, I believe, when we turn to instrumental reasoning and the instrumental principle: a similar defect appears, but it cannot be removed in the same way. Recall Broome's formulation of the instrumental principle:

You should [if you intend to do *E* and believe that doing *M* is a necessary means to doing *E*, intend to do *M*].

This claim is supposed to employ the all-in practical "should," the one that also figures in the principle of means-end transmission:

*Transmission:* If you should do *E*, all things considered, and doing *M* is a necessary means to doing *E*, you should do *M*, all things considered, too.

As before, the interaction of these principles is problematic.

In the original case of *akrasia*, I intend to smoke and believe that buying cigarettes is a necessary means to smoking; but it is not the case that I should buy cigarettes, all things considered. Consider a variation on this case in which there is nothing I can do to change my intention to smoke or my belief about the necessary means: these attitudes are not under my control. It follows that the only way in which I can conform to the conditional [if I intend to smoke and believe buying cigarettes is a necessary means to smoking, then I intend to buy cigarettes] is by intending to buy cigarettes. In other words, intending to buy cigarettes is a necessary means to the truth of that conditional. According to Broome, I should make that conditional true, all things considered. It

25. In "Why Be Rational?" (*Mind* 114 [2005]: 509–63), Niko Kolodny objects that this sort of view omits the normativity of "rational requirements." It depicts the standard of normative practical reasoning as being merely "evaluative" (Kolodny, "Why Be Rational?" 551–55). This gap can be filled by acknowledging that failure to conform to this standard is typically irrational, in a distinctive sense that "ascribe[s] a certain kind of blame" (Stephen L. White, "Rationality, Responsibility and Pathological Indifference," in Flanagan and Rorty, eds., *Identity, Character and Morality*, 401–26, 412). Irrationality in the narrow sense stands to defects of reason as moral culpability stands to moral wrongdoing; it is circumscribed by our capacities. To say that someone is irrational, in this sense, is to ascribe to them a failure of reason they could legitimately have been expected to avoid. (I defend this claim in connection with practical irrationality in "Against Internalism," *Noûs* 38 [2004]: 266–98, secs. 2 and 3.) Failures of normative practical reasoning tend to be irrational in the narrow sense, because they are failures that the agent herself can always recognize as such; other things being equal, they are failures that she is in a position to avoid. This way of understanding the distinctive normativity of "rational requirements"—i.e., of the accusation that someone is being irrational—has the advantage of generality over the proposal with which Kolodny's paper ends ("Why Be Rational?" 557–60). As he acknowledges, his "Transparency" account only applies to normative practical reasoning, while the charge of irrationality is more widespread.

follows, by *Transmission*, that I should intend to buy cigarettes, all things considered. But this need not be so: it may still be false that I should make that decision. It is not vindicated by the fact that I cannot affect the relevant intention (to smoke) or the corresponding means-end belief.

Like the previous objection, to Broome's account of normative practical reasoning, this may seem too quick. But again, I think it is sound. So long as there is no decision that would affect my intention to smoke, or my belief about the means to smoking, there is nothing I can do to change these attitudes, and intending to buy cigarettes will be a necessary means to the truth of Broome's conditional.<sup>26</sup>

The case in question is, admittedly, more peculiar than the one in which I cannot alter my belief that I should  $\phi$ . But it remains possible. Suppose, for instance, that my intention to smoke is sufficiently robust that even if I decided not to smoke, the resulting conflict of intentions would be resolved in its favor; I would still intend to smoke. Nor can I do anything about the belief that buying cigarettes is a necessary means. (In any case, attempting to modify that belief seems like the wrong way to conform to the demands of instrumental reason.) None of this shows that I should buy cigarettes, all things considered.

When we looked at normative practical reasoning, we came upon a parallel problem, but we also found a solution: an appeal to the "objective 'should'" of practical rationality or good practical thought. What happens when we make the same move here?

The most plausible view is that it is part of being practically rational to [give up one's intention to do  $E$  or adopt the intention to do  $M$ ] if one believes that doing  $M$  is a necessary means to doing  $E$ . The disposition to conform to that disjunction when one has the relevant belief is a good disposition of practical thought.<sup>27</sup> The direction of the disposition, running from belief to intention, explains why attempting to modify one's means-end beliefs is the wrong way to conform to the demands of instrumental reason. The disjunctive character of the output of the disposition—roughly: don't intend  $E$  or do intend  $M$ —echoes the wide-scope character of Broome's account. And the fact that it is a

26. Here I employ the modest interpretation of "something I could do" that figured in the earlier objection to Broome: what I can do is what I would do if I made some appropriate decision. As I said in the note attached to that discussion, adopting a more stringent interpretation of what an agent "can do" (for instance, one on which I must be able to make the relevant decision) would only make it easier to find a case in which I cannot do anything about the antecedent of Broome's conditionals, so that making true the consequents is a necessary means to making the whole conditionals true.

27. Compare Broome ("Does Rationality Give Us Reasons?" 322) on the "rational requirement" of means-end reason, which replaces the "objective 'should'" of his earlier account.

disposition actually to conform to that disjunction, not just to be moved in that direction to some degree, captures the peremptory character of the instrumental principle: its insistence on what one must do.

Unfortunately, this view has the same defective implication as Broome's. We can see this if we return to the close connection between reasons for action and good dispositions of practical thought:

*Reasons:* The fact that  $p$  is a reason for A to  $\phi$  just in case A has a collection of psychological states, C, such that the disposition to be moved to  $\phi$  by C-and-the-belief-that- $p$  is a good disposition of practical thought, and C contains no false beliefs.

Together with the view proposed in the previous paragraph, this principle implies that the fact that doing  $M$  is a necessary means to doing  $E$  is always a reason to [give up one's intention to do  $E$  or adopt the intention to do  $M$ ]. This is so far consistent with the example given above. But a conflict is lurking. Corresponding to *Reasons* is the claim that a given fact, that  $p$ , is a *decisive* reason for A to  $\phi$  just in case A has a collection of psychological states, C, such that it is part of good practical thought to be disposed to  $\phi$  (not simply to be moved to  $\phi$  to some degree) if one has C-and-the-belief-that- $p$ , and C contains no false beliefs. It follows from this, together with the present interpretation of the instrumental principle, that the fact that doing  $M$  is a necessary means to doing  $E$  is always a decisive reason to [give up one's intention to do  $E$  or adopt the intention to do  $M$ ]. In other words, whenever this fact obtains, one should [give up one's intention to do  $E$  or adopt the intention to do  $M$ ], all things considered. But this consequence is evidently subject to the same objection as Broome's account. If buying cigarettes is a necessary means to smoking, it implies that I should [give up my intention to smoke or adopt the intention to buy cigarettes], all things considered. If there is nothing I can do to alter my intention to smoke, intending to buy cigarettes will be a necessary means to the truth of that disjunction. It follows by *Transmission* that I should intend to buy cigarettes, all things considered. Once again, however, this is not the case: my decision to buy cigarettes is not justified by the fact that I cannot get rid of my intention to smoke.

What blocked a similar resurgence of problems for normative practical reasoning was the fact that the good disposition there was triggered, in the troublesome cases, by a false belief (that there was a reason to  $\phi$ ) and therefore did not generate a reason. That is why we can accept the "subjective 'should'" account of why one should be moved by one's beliefs about what there is reason to do. By contrast, we cannot accept the "subjective 'should'" account of instrumental reason, on which it is a good disposition of practical thought to [give up one's intention to do  $E$  or adopt the intention to do  $M$ ] when one believes that doing  $M$

is a necessary means to doing *E*. For, in the problem cases, this disposition is triggered by a true belief, and therefore *would* generate a reason, and, moreover, a decisive one. It thus supports an all-in practical “should” quite similar to the one that figures in Broome’s account. At any rate, it is similar enough to be subject to the same objection.

The moral of these arguments is that, surprising though it may seem, instrumental reason is not a dimension of practical reason, after all. The “should” that figures in the instrumental principle is not the all-in practical “should”—applied straight, or in Broome’s distinctive way. Nor can it be the “should” of practical rationality or good practical thought. But if it belongs to practical reason, what else could it be? What other interpretation can we give to a distinctively practical, and rational, “should”?<sup>28</sup>

It may seem that we are entitled to a more radical conclusion: that there is no satisfactory interpretation of the instrumental principle.<sup>29</sup> In what follows, I argue that things are not so bad. We can make sense of the instrumental principle as a principle of theoretical reason, one that employs the epistemic “should.”

### III

There are precedents for this “cognitivist” approach in the work of Gilbert Harman and Jay Wallace.<sup>30</sup> They claim, in different ways, that the instrumental principle can be conceived as the application of theoretical reason to the beliefs that figure in our intentions.

My defense of cognitivism will rest on a controversial account of what an intention is. I have argued for this account elsewhere.<sup>31</sup> Here, I simply assume that it is true and provide a sketch of the necessary details. Although the account is controversial, its basic outlines are familiar enough: it is a development of the claim that intention involves belief. As Anscombe pointed out, the verbal expression of one’s inten-

28. A final proposal: we “should” conform to the instrumental principle in that an ideal practical thinker would conform to it, so that a failure to do so always indicates a defect of practical reason. (Compare the “ideal-world” interpretation of “should” or “ought” familiar to deontic logicians.) The problem with this interpretation is that it does not follow from the fact that an ideal thinker would do something that I should do it, in the circumstance in which I find myself. We need to capture the force of the instrumental principle as it applies to agents in nonideal conditions; and when we do so, we fall into the problems described above.

29. In particular, it can be hard to see how the instrumental principle could fail to count as part of practical reason, since, however we formulate the principle, it must be concerned with what we should intend. I return to this objection, and to the corresponding picture of practical reason, in Sec. IV.

30. Harman, “Practical Reasoning,” 152–53; Wallace, “Normativity, Commitment and Instrumental Reason,” sec. 4.

31. In pt. 1 of *Reasons without Rationalism*.

tion to  $\phi$  is the assertoric utterance of the sentence “I am going to  $\phi$ ,” and thus the expression of belief that one is going to  $\phi$ ; one cannot intend to do something without having that belief.<sup>32</sup>

More carefully, the attitude of intending to do something is a matter of motivating or desire-like belief. Intention represents its object as true in the same way that belief does; under the right conditions, it will constitute knowledge. But it also motivates action after the fashion of desire. The intention to do something has the power to cause one to do it, if one can—sometimes directly, sometimes through taking the necessary means. Intending to  $\phi$  is roughly a matter of having the desire-like or motivating belief that one is going to  $\phi$ .<sup>33</sup>

I say “roughly” because there is a final complication: intention is self-referential. When I intend to  $\phi$ , the content of my attitude is that I am going to  $\phi$  because of that very intention: intention represents itself as motivating action.<sup>34</sup> This claim may seem peculiar. Again, it is something I have argued for elsewhere, and I won’t repeat those arguments here.<sup>35</sup> But the suggestion is not as strange as it sounds. It is part of what one believes in deciding to do something that one’s choice will be efficacious; without that belief, decision would make no sense. The doctrine of self-reference builds this into the content of one’s intention. In doing so, it explains why one cannot intend to do something unintentionally. I can’t intend to drop my keys accidentally, or inadvertently, because intending to drop them is intending to do so in execution of that very intention.

32. G. E. M. Anscombe, *Intention*, 2nd ed. (Oxford: Blackwell, 1963), 1. For the doctrine that intending to  $\phi$  involves the belief that one is going to  $\phi$ , see, esp., Stuart Hampshire and H. L. A. Hart, “Decision, Intention and Uncertainty,” *Mind* 67 (1958): 1–12; Harman, “Practical Reasoning”; and J. David Velleman, *Practical Reflection* (Princeton, NJ: Princeton University Press, 1989), chap. 4. Critics of the doctrine include Donald Davidson, “Intending,” reprinted in his *Essays on Actions and Events* (Oxford: Oxford University Press, 1984), 83–102, 91–94; Michael Bratman, “Intention and Means-End Reasoning,” *Philosophical Review* 90 (1981): 252–65; and Alfred Mele, *Springs of Action* (Oxford: Oxford University Press, 1992), chap. 8.

33. In “The Humean Theory of Motivation” (*Mind* 96 [1987]: 36–61, 54–56), Michael Smith argues that an attitude cannot be both belief-like and desire-like with respect to the same proposition. I dispute his argument in *Reasons without Rationalism*, 49–51.

34. For versions of this claim, see Harman, “Practical Reasoning,” sec. II, and *Change in View* (Cambridge, MA: MIT Press, 1986), 85–86; John Searle, *Intentionality* (Cambridge: Cambridge University Press, 1983), 83–90; and Velleman, *Practical Reflection*, 88–90, 94–97, 140–41. Harman (“Practical Reasoning,” sec. 2, and *Change in View*, 80–81) distinguishes “positive” and “negative” intentions, only the former of which present themselves as causes of action. But his “negative” intentions are causes, too. It is just that the action they cause is overdetermined: it would happen without them. So long as we reject, or qualify, the counterfactual test for causation, we can claim that intentions always satisfy the formula given in the text.

35. See *Reasons without Rationalism*, 41–45.



It is because intention involves belief that intentional action involves what Anscombe called “knowledge without observation.”<sup>36</sup> In the paradigm case of intentional action, I know what I am doing, and I know this spontaneously, in a way that does not turn on observational evidence, though it may depend on background knowledge of the circumstance.<sup>37</sup> More strongly, and more carefully, my knowledge seems to be noninferential, not based on an inference from sufficient prior evidence. It is not a theoretical conclusion—but it is knowledge all the same. This is puzzling in two ways. First, why should motivation, even motivation by reasons, have to go through knowledge of what one is doing?<sup>38</sup> Second, how can this be knowledge, if it is not based on evidence? The answer to the first question, at least in part, is that in the paradigm case of doing something intentionally, I intend to be doing it—here I shift from prospective to progressive intention, or intention in acting—and thus believe that I am doing it. This belief, being part of my intention, derives from practical thought, not theoretical evidence. As well as being partial, this answer makes the second question more acute: how is knowledge in intention possible on the basis of practical thought? That is a topic for another occasion.<sup>39</sup> What we need at present is the fact of its possibility. So long as one knows *how* to act in a certain way and that one will not be prevented, one can form the corresponding intention, and the belief that it involves, without needing evidence and without any violation of theoretical reason. In this respect, intention is quite different from ordinary belief. As I have said, it is a puzzle how this difference is possible. But it must be, if the will is not an illusion, and so for present purposes, I set the puzzle aside.

The possibility of forming beliefs without evidence by forming intentions is the key to understanding the instrumental principle as part of theoretical reason. It follows from the principle of closure, according to which we should believe the logical consequences of our beliefs. This epistemic “should” takes wide scope. Thus:

*Closure:* You should [if you believe that *p* and believe that if *p*, *q*, believe that *q*].

36. Anscombe, *Intention*, 15.

37. Why only the “paradigm case”? Because it is possible to act intentionally without knowledge, as when one tries to do something without being sure one can—and one happens to succeed. For cases of this kind, see Davidson, “Intending,” 91–94. In *Reasons without Rationalism*, 24–26, I show how Anscombe’s claim can be revised so as to avoid the problem.

38. This question animates the exploration of intentional action in pt. 1 of *Reasons without Rationalism*.

39. I have tried to address it in “Practical Knowledge” (unpublished manuscript, University of Pittsburgh, 2007).

It is not that, whenever you believe that  $p$  and believe that if  $p$ ,  $q$ , you should believe that  $q$ , *simpliciter*. Instead, you can satisfy *Closure* in any of three ways: by ceasing to believe that  $p$ , or that if  $p$ ,  $q$ , or by coming to believe that  $q$ .

*Closure* must be thought of as a principle of “ideal rationality.” It is a condition to which we should conform, whatever the circumstance; but it would be misleading to say that everyone who violates this principle is being irrational. As I have argued elsewhere, the charge of “irrationality” can be heard in a narrow way that “ascribe[s] a certain kind of blame.”<sup>40</sup> Irrationality in the narrow sense stands to defects of reason as moral culpability stands to moral wrongdoing; it is circumscribed by our capacities. To say that someone is irrational, in this sense, is to ascribe to them some failure of reason that they could legitimately be expected to avoid. It is unrealistic to expect us to keep track of all our beliefs, and their consequences, in the way that *Closure* claims we should. But this is just to say that we are not always irrational, in the narrow sense, when we violate the principle; it is not an objection to *Closure* as it is meant to be understood.<sup>41</sup>

Even as a principle of ideal rationality, *Closure* might be denied, but not in ways that undermine its employment here. If one believes that there is epistemic (not just practical) reason to avoid cluttering one’s mind with trivialities, one will restrict the demand for closure to beliefs that concern matters of practical significance.<sup>42</sup> Since that condition is met by the beliefs that are relevant to the instrumental principle, the restriction can safely be ignored. In a similar way, reflection on the so-called “paradox of the preface” may prompt a rejection of *Closure* that involves replacing it with a principle of probabilistic coherence for degrees of belief.<sup>43</sup> (Where the degree is one, ideal rationality would still demand closure under logical consequence.) Though it would complicate the treatment of instrumental reason to proceed in these terms, I don’t see why it would be impossible. The result would be a theory of greater generality than the one that I propose, which might then be conceived as a rough approximation, or a limiting case. For simplicity, I will continue to operate with *Closure* itself.

40. White, “Rationality, Responsibility and Pathological Indifference,” 412; Setiya, “Against Internalism,” secs. 2 and 3.

41. Here I agree with Robert Stalnaker, *Inquiry* (Cambridge, MA: MIT Press, 1984), 84; and David Christensen, *Putting Rationality in Its Place* (Oxford: Oxford University Press, 2004), 150–52. It follows that arguments about the difficulty of conforming to certain epistemic standards—as, for instance, in Harman’s *Change in View*—are relevant only to what counts as irrational, in the narrow sense, not to ideal rationality. They are less significant than they might appear.

42. For discussion of this idea, see Harman, *Change in View*, 12–15, 55.

43. This path is pursued in detail by Christensen, in *Putting Logic in Its Place*.

What happens when *Closure* is applied to the beliefs that figure in our intentions? Since intending to  $\phi$  involves the belief that one is going to  $\phi$ , we can derive the following epistemic “should”:

You should [if you intend to do  $E$  and believe that you will do  $E$  only if you do  $M$ , believe that you are going to do  $M$ ].

Again, this principle has wide scope, and one can conform to it in any of three ways.

- (1) You can revise your belief about the relation between doing  $E$  and doing  $M$ .

This is not a violation of instrumental reason, though it may be mistaken on other grounds. It is a kind of wishful thinking to go from intending to do  $E$ , and the belief that you are not going to do  $M$ , to the conclusion that doing  $M$  is not necessary for doing  $E$ .

- (2) You can give up your intention to do  $E$ .
- (3) You can form the belief that you are going to do  $M$ .

One way to form the latter belief is to form the intention of doing  $M$ . If we could show that this is the only epistemically permissible way to form that belief, we would have established, on grounds of theoretical reason, that you should [if you intend to do  $E$ , and believe that you will do  $E$  only if you do  $M$ , intend to do  $M$ ]. We would have established an epistemic instrumental principle.

Before I attempt to fill the gap in this account—why form the intention, not just the belief?—I need to deal with a complication. It turns out that there are circumstances in which it is permissible to form the relevant belief without the corresponding intention—both epistemically permissible, and permitted by instrumental reason, properly understood. We need to refine the conditions in which the instrumental principle applies. Through these refinements we can see when, and why, one must intend the means or give up one’s intention for the end, in order to meet the demands of theoretical reason.

It will be helpful to focus on a specific example. Imagine that I intend to make a fire, and I have an array of beliefs about the conditions that must obtain in order for this to happen. There are cases in which, although I know that I am going to make a fire only if I do  $M$ , it is not a violation of instrumental reason to believe that I am going to do  $M$  without forming the intention to do it. Consider, first, the side effects of making a fire. I know, for instance, that I am going to make a fire only if I make smoke. But I need not intend to make smoke: it is merely a foreseen consequence of what I intend to do. It is quite permissible here to form the belief without the intention. A second case involves what we may call “automatic means”: things I need to do in order to

achieve my end, and which are genuine means to it, but which I need not do intentionally. Thus, I know that I am going to make a fire only if I cause certain chemical reactions in the wood, and if I flex the muscles in my hand so as to light a match for the kindling. But these are means that, so to speak, take care of themselves. I don't have to think about them or care about them; they are parts of a mechanism I can take for granted. So there is no pressure to intend them as means to my end. I may simply conclude that they are going to happen, since I am going to make a fire.<sup>44</sup> Similarly, if I am going to make a fire, I had better not lock myself in the basement and throw away the key. But since I have no inclination to do so, I need not decide against that path.<sup>45</sup> Finally, there is the phenomenon of self-trust, as when I need to perform an intentional action as a means to my end, but I rely on myself to form the relevant intention when the time comes. Again, there is no need to form the intention now: it not a violation of instrumental reason simply to form the corresponding belief. I know that I will need to strike a match in order to light the kindling in order to make a fire. But I am sure that I will decide to so when the materials for the fire are ready and that it is not necessary to settle this ahead of time. A proper formulation of the instrumental principle would deny that I must form the intention to strike a match, instead of the belief that I will—although forming it remains a possibility.

The moral of these cases is that instrumental reason does not generally require that you should [if you intend to do *E*, and believe that you will do *E* only if you do *M*, intend to do *M*]. It is only when you believe a decision is necessary that the instrumental principle gets a grip:<sup>46</sup>

You should [if you intend to do *E* and believe that you will do *E* only if you do-*M*-because-you-now-intend-to-do-*M*, intend to do *M*].

It is because the belief that a decision is necessary is absent when we are dealing with side effects, with automatic means, and with self-trust,

44. On automatic means, see Harman, *Change in View*, 110–11.

45. For Harman, this case would involve a “negative” intention (“Practical Reasoning,” sec. 2, and *Change in View*, 80–81).

46. For formulations of the instrumental principle that accommodate this point, see Harman, “Practical Reasoning”; Wallace, “Normativity, Commitment and Instrumental Reason”; and Bratman, “Intention, Belief, Practical, Theoretical.” It does not follow from the fact that one regards one’s intention as itself a necessary means that forming the intention is an intentional action performed as a means to an end—as though one had an instrumental relation to one’s own will. (Here I disagree with Harman [“Practical Reasoning,” 155–58].) An obvious regress shows that forming an intention cannot always be something one does on the basis of a prior intention. And Gregory Kavka’s “toxin puzzle” arguably shows that it never is (“The Toxin Puzzle,” *Analysis* 43 [1983]: 33–36).

that it is instrumentally permissible to form the belief that you are going to do *M* without forming the intention to do it.

This affects the project of cognitivism about instrumental reason. We need not show that forming the intention to do *M* is the only epistemically permissible way to form the belief that you are going to do *M*, whenever you intend to do *E* and believe that you will do *E* only if you do *M*. This is not the case. You may permissibly conclude that you are going to do *M* on the ground that you are going to do *E*, without yet intending to do *M*, if doing it would be a side effect of doing *E*, or an automatic means, or when you predict that you will decide to do *M* later on. These are not violations of instrumental or of theoretical reason.

Applying *Closure* to the beliefs that figure in our intentions, we can derive the following, more intricate, epistemic “should”:

You should [if you intend to do *E* and believe that you will do *E* only if you do-*M*-because-you-now-intend-to-do-*M*, believe that you are going to do *M* because you now intend to do it].

As with the simple version of cognitivism above, this is a wide-scope “should,” which can be satisfied in three ways: by revising your beliefs about what is necessary for doing *E*—not a violation of instrumental reason, but perhaps a mistake on other grounds—by giving up your intention to do *E*, or by forming the belief that you are going to do *M* because you now intend to do it. One way to form the last belief is to form the intention of doing *M*: to intend to do *M* is, *inter alia*, to believe that one is going to do *M* because of that very intention.

The question is: why should forming the intention be the only permissible way to form that belief? How could theoretical reason prohibit “mere” belief? After all, it is surely possible for an “agent [to] be mistaken about his own intentions.”<sup>47</sup> It is possible to believe that one is going to do *M* because one so intends, even though one does not. The consequent problem for the cognitivist has been developed in ingenious detail by Michael Bratman:<sup>48</sup> “Suppose I believe that I intend to [do *M*] but in fact do not so intend. . . . Still, my belief fills the gap *in my beliefs* which threatens [epistemic] incoherence. . . . But though the demand for [epistemic] coherence is met, the demands of means-end rationality are not. I intend some end, believe that to achieve it I must decide on some means, and have in fact reached no such decision.”<sup>49</sup> Bratman is describing a case in which I satisfy the epistemic

47. Bratman, “Intention and Means-End Reasoning,” 256 n. 4.

48. The development appears in Bratman, “Intention, Belief, Practical, Theoretical”; here I cite an earlier and briefer discussion, from “Intention and Means-End Reasoning.”

49. Bratman, “Intention and Means-End Reasoning,” 256 n. 4.

“should,” above, by forming the belief that I am going to do *M* because I intend to do it, without actually forming that intention. Thus, instead of intending to buy wood, in order to make a fire, when I realize that this intention is necessary, I merely believe that I am going to buy wood because I intend to do so. On the face of it, it is possible in this way to violate the instrumental principle without epistemic defect, and if that is so, the cognitivist account of instrumental reason is false.

In discussing a similar objection to a similar view, Wallace insists that it is “independently irrational for you to have false beliefs about the content of your intentions”: “you will be subject to rational criticism if you believe that you intend to do *y* without really so intending.”<sup>50</sup> If this were true, it would solve the problem—but why accept it? The mind is not transparent to itself. We can have false beliefs about our own emotions and desires, for instance, without transgressing theoretical reason: we can simply make mistakes. Why should things be different when it comes to false beliefs about how we intend to act? Unless we can answer that question, Bratman’s objection will stand.

I am optimistic, though, that an answer can be found, at least on the theory of intention sketched above; we can explain what is defective about the false belief that one intends to  $\phi$ . The explanation turns on the incoherence of a certain kind of inference, one that appears in a related but different case. Consider the following passage from Richard Moran’s *Authority and Estrangement*, about the subject-perspective on one’s own beliefs: “[The] beliefs of another person may represent indicators of the truth, evidence from which I may infer some conclusion about the matter. I may trust or mistrust them. With respect to my own beliefs, on the other hand, there is no distance between them and how the facts present themselves to me, and hence no going from one to the other.”<sup>51</sup> This description seems right. It would be peculiar to infer that *p* on the ground that I believe that *p*. But why? One reason is that an inference of this kind could never be both sound and ampliative. Whenever I would form a new belief by inferring that *p* from the belief that I believe that *p*, the premise of my inference is false: I did not, in fact, believe that *p*. Whenever the premise is true, however, the inference is redundant: I already believe that *p*. In the epistemic sense of “should,” one should never make an inference of this pathological kind. Hence the force of Moran’s remark.

What does this have to do with intention and instrumental reason? It is a curious fact that the belief that one intends to  $\phi$  can only be acquired, on the basis of evidence, by an inference that is defective in

50. Wallace, “Normativity, Commitment and Instrumental Reasoning,” 22.

51. Richard Moran, *Authority and Estrangement* (Princeton, NJ: Princeton University Press, 2001), 75.

just this way. An inference to the conclusion that I intend to  $\phi$ , from *any* premise, will instantiate a pattern that cannot be both sound and ampliative. For suppose that I infer that I intend to  $\phi$  on the ground that  $p$ . If the conclusion is false, the inference is unsound. But if the conclusion is true, the self-reference of intention ensures that the inference is redundant. If I intend to  $\phi$ , I already believe that I am going to  $\phi$  because I so intend. Theoretical inference cannot support the belief that I intend to  $\phi$  without involving a defect that amounts to incoherence—like inferring the facts from my own beliefs.

It follows, I think, that there is something incoherent about the belief that I intend to  $\phi$ , unless it is constituted by the intention to  $\phi$ . It is an inherently defective belief.<sup>52</sup> If this is right, it closes the gap in the cognitivist account of instrumental reason. In the epistemic sense of “should,” you should believe the logical consequences of your beliefs. Thus you should [if you intend to do  $E$  and believe that you will do  $E$  only if you do- $M$ -because-you-now-intend-to-do- $M$ , believe that you are going to do  $M$  because you now intend to do it]. But the final belief is defective or incoherent, in epistemic terms, except when it takes the form of intending to do  $M$ . It follows that, in the epistemic sense, you should [if you intend to do  $E$  and believe that you will do  $E$  only if you do- $M$ -because-you-now-intend-to-do- $M$ , intend to do  $M$ ]. The instrumental principle belongs to theoretical reason.<sup>53</sup>

#### IV

Think back to the case of *akrasia*, in which I intend to smoke, but should not buy cigarettes, all things considered. I am subject to the epistemic instrumental principle: I should [if I intend to smoke and believe that I will smoke only if I buy-cigarettes-because-I-now-intend-to-do-so, intend to buy cigarettes]. But it does not follow that I should conform to that conditional in the practical sense of “should,” and so the problems from Section II do not arise.

Can we detach a narrow-scope “should” from the epistemic instru-

52. Here I move from dynamic to static epistemology: since it is epistemically permissible to form the belief that one intends to  $\phi$  only by forming that intention, it is an epistemic failing to have that belief when one does not intend to  $\phi$ . This is not to say that it is always irrational to believe that I intend to  $\phi$  when I do not. I can be forgiven for mistaking or misremembering my plans, or for being misled by the evidence about my own intentions. The point is about incoherence and the epistemic “should,” not about culpability.

53. Since it relies solely on the cognitive conditions of intending, this account implies that, whenever you believe you are going to do  $E$ , and believe that you will do  $E$  only if you do- $M$ -because-you-now-intend-to-do- $M$ , you are in same position as when you intend to do  $E$ . This may be surprising, but it is quite correct. In the circumstance described, you must conclude that you are not going to do  $E$ , after all—unless you form the intention to do  $M$  or revise one of the connecting beliefs.

mental principle? Suppose, for instance, that I cannot affect my intention to smoke or my beliefs about the necessary means, so that intending to buy cigarettes is the only way to make the relevant conditional true. Still, it wouldn't follow I should intend to buy cigarettes, all things considered, in the epistemic sense. For the principle of *Transmission* does not apply to the epistemic "should," which directs our thinking, not the pursuit of means to ends. In any case, even if we could derive that epistemic "should" under some condition or other, this would not be problematic: it would not entail that I should buy cigarettes, or intend to do so, in the practical sense.<sup>54</sup>

These arguments assume that there is a distinction to be made between practical and theoretical reason, and they tell us something about the character of that distinction. One way to see this is to examine a paradox that has been hovering in the background since the end of Section II. How can one possibly deny that the instrumental principle is "practical"? A standard way of drawing the distinction between practical and theoretical reason is to say that the former is concerned with what to intend, the latter with what to believe; and the instrumental principle is certainly concerned with what we should intend. If it is an epistemic principle, one might think, it is also a principle of practical reason.<sup>55</sup> If I am right, however, this must be a mistake. We can show directly, from assumptions about the all-in practical "should," and its relation to practical rationality or good practical thought, that the instrumental principle does not belong to practical reason. (That was the argument of Sec. II.) It follows that the standard characterization of practical reason is not correct; it is, at any rate, simplistic. A better account would add that good practical thought, unlike theoretical reason, is indifferent to the truth of our beliefs.<sup>56</sup> In this sense, practical

54. How would such a derivation go? It would have to take us from a case in which I should believe that I am going-to-do-*M* because I now intend to do it, by way of the principle (from the end of Sec. III), that I should [believe that I intend to do *M* only if I have that intention], to the conclusion that I should intend to do *M*, in the epistemic sense. This argument has the following invalid form: I should believe that I  $\phi$ ; I should [believe that I  $\phi$  only if I  $\phi$ ]; so, I should  $\phi$ . This must be invalid because the final "should" will not make sense, as epistemic, when doing or being  $\phi$  is not a matter of my cognitive state. That this condition fails in the present case is masked by the fact that intending to do *M* partly consists in the belief that one is going to do *M*. But it also consists in a motivational condition that theoretical reason cannot govern. So, it won't make sense to claim that I should intend to do *M*, in the epistemic sense of "should."

55. In defending versions of cognitivism about the instrumental principle, Harman ("Practical Reasoning") and Wallace ("Normativity, Commitment and Instrumental Reason") do not deny that their topic is practical reason.

56. This is arguably so even for beliefs about practical reason itself. Reasoning about practical rationality, and about what one should do, is theoretical reasoning; its standards are epistemic. But the connections here are complicated.



reason is nontheoretical. Thus the conclusion of Section II leaves room for cognitivism about instrumental reason.<sup>57</sup>

This leads to a deeper point, about the significance of cognitivism for practical philosophy. The appeal of the instrumental principle is a constant temptation to believe that practical reason must be, at some level, indifferent to morality. It is a source of pressure behind the question, “Why should I be just or benevolent, when that conflicts with my ends?” If we can understand means-end reason, in general—not just the special case of necessary means—as part of theoretical rather than practical reason, this source of pressure will subside.<sup>58</sup> There may be other grounds on which to claim that practical reason is amoral, but the fact that one should always take the means to one’s ends will no longer be among them.

57. It does not leave room for the global cognitivism apparently found in Velleman’s *Practical Reflection*. On this view, practical reason can be identified with theoretical reason, applied to our intentions. (For this reading of Velleman, see Bratman, “Cognitivism about Practical Reason.”) If the instrumental principle is theoretical, but not practical, this identification must be false.

58. The conditional here is not rhetorical: I have only dealt with the special case. In doing so, I appealed to *Closure* as a standard for full belief. Despite Broome’s skepticism (“Practical Reasoning,” 109), my hope is that we can begin to explain the aspects of instrumental reason that govern nonnecessary means by turning to the probabilistic replacement for *Closure* that we most likely need in any case.