



---

The Possibility of Practical Reason

Author(s): J. David Velleman

Reviewed work(s):

Source: *Ethics*, Vol. 106, No. 4 (Jul., 1996), pp. 694-726

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/2382031>

Accessed: 03/01/2012 10:44

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

# The Possibility of Practical Reason\*

*J. David Velleman*

Suppose that reasons for someone to do something must be considerations that would sway him toward doing it if he entertained them rationally.<sup>1</sup> And suppose that the only considerations capable of swaying someone toward an action are those which represent it as a way of attaining something he wants, or would want once apprised of its attainability.<sup>2</sup> These assumptions, taken together, seem to imply that

\* I am grateful to Sonja Al-Sofi, Stephen Darwall, Jennifer Church, Paul Boghossian, Alfred Mele, Elijah Millgram, Derek Parfit, Peter Railton, Sigrun Svavarsdottir, Nicholas White, Bernard Williams, Stephen Yablo, and several anonymous referees for helpful discussions of the issues raised in this article. Earlier versions of the paper were presented to the philosophy departments at New York University, Stanford University, the University of Illinois at Urbana-Champaign, and the University of Houston; and to the philosophy faculty colloquium at the University of Michigan. Some of the material was also presented to Michael Bratman's 1993 National Endowment for the Humanities seminar on intention. I have benefited from comments received on all of these occasions, especially the comments of Frances Kamm, Fred Dretske, Rachel Cohon, Allan Gibbard, Sally Haslanger, David Hills, Tomis Kapitan, Jeff McMahan, Patrick Hays, and David Phillips. Finally, I received many helpful comments from participants in a graduate seminar taught at the University of Michigan in the winter of 1993. Work on this article was supported by the Edna Balz Lacy Faculty Fellowship at the Institute for the Humanities, University of Michigan.

1. This principle is meant to apply only to complete sets of reasons, not to reasons taken individually. That is, if a particular consideration counts as a reason only in the context of a larger set or series of considerations, then it need not be capable of swaying the agent unless it is considered in that context. The assumption that 'reasons' denotes complete sets of reasons will be in force throughout the following discussion. I shall also rely on the success-grammar of the word 'considerations': considerations are, by implication, true considerations—or, as I shall sometimes call them, facts.

2. The last clause is meant to account for cases like this: "It may be true of me that were the aroma of fresh apple pie to waft past my nose I would be moved to discover its source and perhaps to try to wangle a piece. It does not follow from this, however, that before I smell the pie I desire to eat it or to eat anything at all" (Stephen L. Darwall, *Impartial Reason* [Ithaca, N.Y.: Cornell University Press, 1983], p. 40). Here apple pie is something that the agent doesn't yet want but will want once he considers its attainability, and so considerations about how to obtain it are capable of influencing

the only considerations that can qualify as reasons for someone to act are considerations appealing to his antecedent inclinations<sup>3</sup>—that is, his desires or dispositions to desire.<sup>4</sup>

This conclusion amounts to an admission that reason really is, as Hume put it, the slave of the passions,<sup>5</sup> and Hume's conclusion is one that many philosophers hope to avoid. Some try to avoid the conclusion by rejecting one of the premises from which it appears to follow.<sup>6</sup> Others prefer to keep the premises while arguing that the conclusion doesn't actually follow from them.<sup>7</sup>

In my view, the question whether reasons do or do not depend on an agent's inclinations should simply be rejected, because it embodies a false dichotomy. This dichotomy has recently come to be formulated in terms introduced by Bernard Williams.<sup>8</sup> In Williams's terminology,

---

him in the requisite way. I take it that this mechanism is what David Hume regarded as the first of the two ways in which reason can influence action: "Reason . . . can have an influence on our conduct only after two ways: Either when it excites a passion by informing us of the existence of something which is a proper object of it; or when it discovers the connexion of causes and effects, so as to afford us means of exerting any passion" (*A Treatise of Human Nature*, ed. by L. A. Selby-Bigge [Oxford: Clarendon, 1978], p. 459).

3. Note that I am not using the word 'inclination' in its Kantian sense. I am using it as the generic term for conative or motivational states of all kinds.

4. The argument presented here is discussed at length in Darwall, *Impartial Reason*, esp. chaps. 2 and 5.

5. Hume, p. 415. I do not claim that the argument offered above for Hume's conclusion should necessarily be attributed to Hume.

6. The most frequent target has been the second assumption, which is sometimes called the Humean theory of motivation. See, e.g., Thomas Nagel, *The Possibility of Altruism* (Princeton, N.J.: Princeton University Press, 1970); John McDowell, "Are Moral Requirements Hypothetical Imperatives?" in *Proceedings of the Aristotelian Society*, suppl. ser., 52 (1978): 13–29; Darwall, *Impartial Reason*, chap. 5; Rachel Cohon, "Are External Reasons Impossible?" *Ethics* 96 (1986): 545–56. For arguments defending this assumption, see Michael Smith, "The Humean Theory of Motivation," *Mind* 96 (1987): 36–61; and Alfred Mele, "Motivational Internalism: The Powers and Limits of Practical Reasoning," *Philosophia* 19 (1989): 417–36. Arguments against the first assumption are rare, although some philosophers have argued against a related assumption applied to moral requirements rather than reasons for acting. See, e.g., William Frankena, "Obligation and Motivation in Recent Moral Philosophy," in *Essays on Moral Philosophy*, ed. A. I. Melden (Seattle: University of Washington Press, 1958), pp. 40–81; and David Brink, *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press, 1989), chap. 3.

7. See Christine Korsgaard, "Skepticism about Practical Reason," *Journal of Philosophy* 83 (1986): 5–25.

8. Bernard Williams, "Internal and External Reasons," in *Moral Luck* (Cambridge: Cambridge University Press, 1981), pp. 101–13; "Internal Reasons and the Obscurity of Blame," in *Making Sense of Humanity and Other Philosophical Papers* (Cambridge: Cambridge University Press, 1995), pp. 35–45; "Replies," in *World, Mind, and Ethics; Essays on the Ethical Philosophy of Bernard Williams*, ed. J. E. J. Altham and Ross Harrison (Cambridge: Cambridge University Press, 1995), pp. 185–224.

“internal” reasons are those which count as reasons for someone only by virtue of his antecedent inclinations; “external” reasons are those which count as reasons for someone independently of his inclinations.<sup>9</sup> The Humean conclusion implies that all reasons are internal, in this sense, and it is therefore called internalism; its denial is called externalism.<sup>10</sup> My thesis is that we do not in fact have to choose between the two.

\* \* \*

Christine Korsgaard has pointed out that the foregoing argument doesn't necessarily yield any constraint on what counts as a reason for acting.<sup>11</sup> It may instead yield a constraint on who counts as a rational agent.

9. Here I am choosing one of two possible readings that have occasioned considerable confusion in the literature. The confusion can be traced to the casual manner in which Williams introduces the term ‘internal reason’. Williams carefully defines what he calls the “internal interpretation” of the statement “A has reason to  $\phi$ .” Interpreted internally, the statement implies that A has some motive that can be served by his  $\phi$ -ing. Williams then says “I shall also for convenience refer sometimes to ‘internal reasons’ and ‘external reasons’” (“Internal and External Reasons,” p. 101). But Williams never explains how a scheme for interpreting reason-attributions can be transformed into a scheme for classifying reasons themselves.

Two different schemes of classification have suggested themselves to philosophers writing in this area. One scheme classifies as internal any reason that can engage one of the agent's motives so as to sway him toward doing that for which it is a reason. The other scheme classifies as internal only those reasons whose status as reasons depends on their capacity to engage the agent's motives in this way. An internal reason, on this latter scheme, is one that wouldn't be a reason if the agent didn't have a motive that it could engage. The difference between these schemes of classification can be illustrated by the case of an agent who has both a reason and a corresponding motive. According to the first scheme, this reason is definitely internal, since the agent has a motive corresponding to it. According to the second scheme, however, this reason could still be external, if it would remain a reason for the agent whether or not he had the motive.

Only the latter scheme captures the entailment that distinguishes Williams's “internal interpretation” of reason-attributions. For on the former scheme, the agent's having a reason doesn't require him to have a motive. If he lacks a corresponding motive, then his reason doesn't necessarily cease to be a reason, on this scheme; it simply ceases to be internal. Yet under the internal interpretation of reason-attributions, the agent must have the motive in order for it to be true that he has a reason at all. I therefore prefer the latter scheme of classification.

10. Note that this usage differs somewhat from that of other philosophers, for whom the term ‘internalism’ refers to our first premise, requiring reasons to have the capacity of exerting an influence.

11. Korsgaard, “Skepticism about Practical Reason.” For other discussions of Williams, see Cohon, “Are External Reasons Impossible?” and Rachel Cohon, “Internalism about Reasons for Action,” *Pacific Philosophical Quarterly* 74 (1993): 265–88; Martin Hollis, *The Cunning of Reason* (Cambridge: Cambridge University Press, 1987), chap. 6; Brad Hooker, “Williams' Argument against External Reasons,” *Analysis* 47 (1987): 42–44; John McDowell, “Might There Be External Reasons?” in Altham and Harrison, eds., pp. 68–85; Elijah Millgram, “Williams' Argument against External Reasons,” *Nous* 30 (1996): 197–220.

The first premise of our argument doesn't entail that if a consideration fails to influence someone, then it isn't a reason for him to act; it entails that if a consideration fails to influence someone, then either it isn't a reason for him to act or he hasn't entertained it rationally. The inclinations that would make an agent susceptible to the influence of some consideration may therefore be necessary—not to the consideration's being a reason for him—but rather to his being rational in entertaining that reason. And our premises may consequently imply that an agent's inclinations determine, not what he has reason for doing, but whether he is rational in his response to the reasons he has.

Korsgaard favors the latter conclusion over the former.<sup>12</sup> In denying the dependence of reasons on inclinations, she qualifies as an externalist, in Williams's terminology.<sup>13</sup>

Korsgaard's critique of Williams suggests a version of externalism that goes something like this.<sup>14</sup> Being a rational agent entails having

12. See also Michael Smith, "Reason and Desire," *Proceedings of the Aristotelian Society* 88 (1988): 243–58. Smith seems to think (pp. 248–52) that he and Korsgaard disagree, but I think that they don't. In particular, Smith believes that Korsgaard rejects the second premise, that considerations can influence an agent only in conjunction with his conative attitudes. But I don't interpret Korsgaard as rejecting this premise. Korsgaard never claims that a consideration, or belief, can move an agent without the help of a conation or motive; what she claims, I think, is that the desires and values mediating the influence of a consideration need not be ordinary motives, of the sort that are directed at the agent's ends, since they can instead constitute his virtue of rationality.

Williams responds to this argument but seems to misunderstand it. He seems to think that if all rational agents have, say, a motive for doing what's right, then the fact that an action is right will turn out to be an internal reason for them, after all: "If this is so, then the constraints of morality are part of everybody's [motivational set], and every correct moral reason *will be* an internal reason" ("Internal Reasons and the Obscurity of Blame," p. 37). But here Williams adopts a sense of the phrase 'internal reason' that fails to capture his own "internal interpretation" of reason-attributions, as I have explained in n. 9, above. In this sense, an internal reason is one with the capacity to engage an agent's motives, but not necessarily one whose very status as a reason depends upon that capacity.

I believe that Korsgaard is working with the alternative (and, to my mind, preferable) sense of 'internal reason', according to which an internal reason is one whose status as a reason depends on its capacity to engage the agent's motives. And what Korsgaard envisions is that reasons for behaving morally will qualify as reasons whether or not people have motives that such reasons can engage. Even if people happen to have the relevant motives, reasons for behaving morally will still be independent of them, in Korsgaard's view, and such reasons should be classified as external.

For a misinterpretation similar to Williams's, see John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), p. 85, n. 33.

13. I do not mean that Korsgaard would call herself an externalist, since she uses the term in a somewhat different sense. See n. 10 above.

14. I don't mean to claim that Korsgaard holds this version of externalism. Korsgaard's "Skepticism about Practical Reason" seems designed to be independent, in many respects, of her larger metaethical project. It therefore leaves open various versions of externalism that Korsgaard herself would not necessarily endorse. Indeed, I suspect that the version of externalism discussed here in the text corresponds to what Korsgaard

various motives, including a preference for one's own greater good<sup>15</sup> and an acceptance of moral principles.<sup>16</sup> A rational agent is influenced by a reason for doing something when, for example, he considers some respect in which doing it is morally required; and this consideration can influence him because an inclination to abide by moral requirements is partly constitutive of his rationality. If an agent lacks this inclination, its absence won't prevent him from having moral reasons for acting: moral requirements will still count as reasons for him to act. Rather, lacking an inclination to abide by moral requirements will render the agent irrational, by making him insensitive to this particular kind of reason.

### THE EXTERNALIST'S BURDEN OF JUSTIFICATION

One liability of this model is that it must identify particular features of an action as constitutive of reasons for taking it, whether an agent cares about them or not, and it must then criticize an agent as irrational if he should fail to care about those features. The model thus incorporates specific normative judgments, to the effect that one ought to be inclined toward courses of action with the features in question.

What entitles the externalist to build these normative judgments into his model of practical reason? As Williams puts it, "Someone who claims the constraints of morality are themselves built into the notion of what it is to be a rational deliberator cannot get that conclusion for nothing."<sup>17</sup>

Korsgaard does not try to get this conclusion for nothing, however. On the contrary, she insists that the normative judgments built into her conception of practical reason will require an "ultimate justification," which the externalist hopes to provide.<sup>18</sup> Indeed, the possibility of such a justification is the centerpiece of her paper.

Korsgaard's quarrel with Williams, after all, is that he prematurely discounts the possibility of justifying externalism. In assuming that an agent's imperviousness to a consideration impugns its status as a reason, rather than the agent's rationality, Williams assumes that its status as a reason cannot be established independently. For if a consideration could be certified as a reason for someone irrespective of whether he's susceptible to it, then his lack of susceptibility would thereby come to impugn his rationality instead. Yet certifying something as a reason

---

rejects under the label "dogmatic rationalism" in "The Normativity of Instrumental Reason," in *Value and Practical Reason*, ed. Garrett Cullity and Berys Gaut (Oxford: Oxford University Press, in press).

15. Korsgaard, "Skepticism about Practical Reason," p. 18.

16. *Ibid.*, p. 22.

17. Williams, "Internal Reasons and the Obscurity of Blame," p. 37.

18. Korsgaard, "Skepticism about Practical Reason," p. 22.

for someone irrespective of his susceptibilities would amount to showing that it is an external reason, and hence that externalism is true. When Williams presupposes the impossibility of such a showing, he is presupposing the impossibility of justifying externalism. His case for internalism thus rests on antecedent skepticism about the alternative.

So Korsgaard argues—cogently, I believe. Yet even if she is right that the case for internalism rests on skepticism about externalism, the question remains whether we aren't entitled to be skeptical. What are the prospects for showing that something is a reason for someone whether or not he has the inclinations to which it would appeal? How will the externalist demonstrate that there are considerations by which any agent ought to be moved?

\* \* \*

One might think that an externalist could avoid this burden of justification by avoiding the identification of any particular considerations as reasons, or of any particular inclinations as rational. But I doubt whether this strategy can work.

The version of externalism outlined above incurs a burden of justification because it judges an agent to be irrational unless he is inclined to be swayed by particular, substantive features of actions, whose value or importance may be open to question. All that externalism needs to say, however, is that the inclination responsible for the influence of reasons is one that's essential for the agent's rationality. Does this inclination have to be an inclination to be swayed by particular considerations, specified by their substance? Maybe it can be an inclination to do whatever is supported by reasons as such, or whatever is rational as such.

The inclination that's now being proposed isn't an inclination to do things with any particular features, other than the feature of being favored by reasons (whatever they may consist in) or the feature of being rational (whatever that is). Hence the claim that this inclination is essential to an agent's rationality doesn't call for any justification. For how could rationality fail to require an inclination to do what's rational, or what's favored by reasons?

\* \* \*

Unfortunately, this version of externalism doesn't ultimately succeed in shedding the burden of justification, since it doesn't avoid the need to specify what counts as a reason or a rational action. To be sure, all it requires of a rational agent is that he be inclined to act in accordance with reasons or rationality as such. But in order for reasons to influence an agent by way of this inclination, he must recognize them as reasons,

or as evidence of rationality, and so he needs some criterion of what counts as a reason or as a rational action. And until such a criterion is supplied, the proposed version of externalism will be nothing but the trivial assertion that rationality is a disposition to be influenced by reasons.

What's needed to save this version of externalism from triviality is a criterion specifying what it is about an action that makes it rational or constitutes a reason for taking it. And this criterion will once again require justification.

At this point, the externalist may attempt to repeat his earlier evasive maneuver. He is committed to the existence of a criterion by which an agent can recognize reasons or rational actions; but is he committed to its being a substantive criterion, which would have to be justified? Maybe an agent can recognize reasons or rational actions by their satisfying the generic concepts of what it is to be a reason or a rational action as such.<sup>19</sup>

Yet this strategy of continually postponing controversy is unlikely to help. Asking the agent to identify a rational action under the guise of rationality as such, or to identify a reason for acting under the guise of a reason as such, would be somewhat like asking him to hunt for something described only as "the quarry," or to play a game with an eye to something described only as "winning." It would be to assign him a task with a formal object but no substantive object—and hence with no object at all.

### THE OBJECT OF PRACTICAL REASONING

The formal object of an enterprise is a goal stated solely in terms of, or in terms that depend on, the very concept of being the object of that enterprise.<sup>20</sup> Thus, for example, winning is the formal object of a competitive game, since "winning" just is the concept of succeeding in competition. Similarly, the formal object of a search or hunt is the quarry, and the formal object of a question is the answer.

Any enterprise that has a formal object must have a substantive object as well—that is, a goal that is not stated solely in terms that

19. I believe that Korsgaard proposes this very strategy ("Skepticism about Practical Reason," pp. 30–31). And I believe that there may be a way—a distinctively Kantian way—of making the strategy work. I discuss this Kantian version of the strategy briefly in n. 25, below. Note, then, that the present argument does not purport to prove that the strategy in question is unworkable. It's meant to justify doubts about the strategy, by showing just how difficult it will be to carry out.

20. I suspect that the argument offered in this section is related to the argument offered by Williams on pp. 109–10 of "Internal and External Reasons." Because I don't fully understand the relevant passage, however, I hesitate to attribute the argument to Williams.

depend on the concept of being the object of that enterprise.<sup>21</sup> In the case of a competitive game, there must be a substantive object of the game, something that constitutes winning but cannot simply consist in winning, so described. A game whose object was specifiable only as “winning” wouldn’t have an object—that is, wouldn’t have any object in particular. And if a game had no particular object, then there would be no such thing as winning it, and so it wouldn’t be a fully constituted competitive game. Similarly, a hunt whose object was specifiable only as “the quarry” wouldn’t be a fully constituted search, and the question “What is the answer?” isn’t by itself a fully constituted question.

\* \* \*

Since practical reasoning is an enterprise at which one can succeed or fail, it must have an object against which success or failure can be measured. What, then, is the object of practical reasoning?

One might suggest that practical reasoning has the object of figuring out what to do, or answering the question “What shall I do?” But this suggestion either misstates the object of practical reasoning or states it in merely formal terms.

The statement that practical reasoning has the object of figuring out what to do may simply mean that it has the object of arriving at something to do or of issuing in an action. So interpreted, however, the statement is mistaken, since issuing in an action—some action or other—is not the object of practical reasoning. Issuing in an action may be what makes reasoning practical, but the object of such reasoning is, not to issue in just any action, but to issue in some actions rather than others.

The object of practical reasoning must therefore be to arrive at a privileged action or an action in some privileged class. And when “figuring out what to do” is interpreted as expressing this object, it turns out to be a merely formal specification, since “what to do,” so interpreted, simply means the correct or privileged thing to do, the thing whose discovery is being attempted. Hence there must be a further, substantive specification of the action or kind of action that practical reasoning aims to identify. A mode of reasoning whose goal was specified solely as “figuring out what to do” would be like a search

21. The distinction between the formal and substantive aims of practical reason is discussed by Derek Parfit in *Reasons and Persons* (Oxford: Clarendon, 1984), pp. 3, 9, 37. As David Gauthier has pointed out (“Rationality and the Rational Aim,” in *Reading Parfit*, ed. Jonathan Dancy [in press]), Parfit is less than clear on the relation between these aims; in particular, Parfit doesn’t appear to believe that the substantive aim of practical reason, as identified by a particular theory, is a specification of the formal aim. Like Gauthier, I prefer to use the phrase ‘substantive aim’ for that which specifies what it is to achieve the formal aim.

whose object was specified solely as “figuring out where to look,” or a question whose object was specified solely as “figuring out how to reply.”

Similar remarks apply to the notion that practical reasoning aims at figuring out the best thing to do.<sup>22</sup> This notion is correct if ‘the best thing to do’ means “the privileged action”—that is, the action that uniquely satisfies the standard of success for this very reasoning. But in that case, it merely expresses the formal object of the enterprise. There can be an enterprise of figuring out the best thing to do, in this sense, only if that enterprise also has a substantive standard of success, just as there can be an enterprise of figuring out the best way to reply only if there is a substantive question, and there can be an enterprise of figuring out the best place to look only if there is a substantive quarry.

Of course, ‘the best thing to do’ might be interpreted, alternatively, as already expressing a substantive value: it might mean, for example, “the action that’s optimistic,” in the sense that it contributes most to the agent’s welfare or to the welfare of everyone. But in that case, the notion that practical reasoning aims to figure out the best thing to do will once again express a value judgment that calls for justification.

\* \* \*

What, then, about rational action or reasons for acting? Can the object of practical reasoning be to identify a rational thing to do, or a thing that one has reason for doing?

The concepts of rational action and reasons for acting are potentially confusing in that they can have both generic and specific uses. If we specify a substantive kind of action as the object of practical reasoning, then we can grant it the honorific “rational,” so that the phrase ‘rational action’ names actions of the specified kind. Similarly, if we specify substantive features that practical reasoning looks for in an action, we can grant the honorific “reasons for acting” to those features. Practical reasoning will then turn out to aim at the rational thing to do, or at what there is reason for doing, but only because ‘rational’ and ‘reasons’ are names for substantive objects.

What cannot be the aim of practical reasoning is rational action merely as such—that is, action conceived as rational in the generic sense, rather than in a sense defined by a specific standard. The generic concept of rational action is just the concept of action that would issue from competent practical reasoning. Until there is something that

22. This notion is, for example, the basis of Donald Davidson’s conception of practical reasoning. See his *Essays on Actions and Events* (Oxford: Clarendon, 1980). The problems mentioned here are discussed further in the text accompanying n. 45 below.

counts as competence in practical reasoning, nothing counts as a rational action in the generic sense. And competence in practical reasoning can be defined only in relation to the object of the enterprise, since competence is a disposition toward success. To be indicative of competent practical reasoning is to be indicative of practical reasoning that's well suited to achieving its object. Defining the object of practical reasoning as action that's rational in this sense would thus be to string definitions in a circle, leaving the object of practical reasoning still undefined. It would be like trying to teach someone a game by telling him that the object was to make a competent showing; whereas what counts as a competent showing always depends on the substantive object of the game.<sup>23</sup>

Similarly, the sole aim of practical reasoning cannot be action supported by reasons merely as such—that is, reasons conceived under the generic concept expressing what it is to be a reason. The generic concept of a reason for acting is the concept of something that warrants or justifies action. And to justify something is to show or indicate it to be just—that is, in accordance with a *jus*, or rule of correctness. Until there is something that constitutes a correct conclusion or a correct inference, there can be nothing that constitutes justifying a conclusion or an inference, and so there can be nothing that constitutes a reason for a conclusion or an inference, in the generic sense. So, too, until there is something that constitutes correctness in actions, or in outcomes of practical reasoning, there can be nothing that satisfies the generic concept of a justification for action, or a justification in practical reasoning; and so there can be nothing that satisfies the generic concept of a reason for acting.

## JUSTIFYING A SUBSTANTIVE CONCEPTION OF REASONS

This argument suggests that the externalist cannot indefinitely postpone giving substantive characterizations of rationality or reasons. The externalist must at some point provide practical reasoning with a substantive standard of success, which will either consist in or give rise to a substantive account of the features that constitute reasons for an action. The externalist will then have to justify his normative judgment that an agent ought to be swayed by consideration of the specified features.<sup>24</sup>

23. Of course, we could introduce a substantive conception of competent practical reasoning—a substantively specified procedure, adherence to which constitutes good reasoning. In that case, however, the definition of rational action as that which would issue from competent practical reasoning will become a substantive concept, which once again requires justification.

24. The notion that practical reasoning is framed by a criterion of success for actions is discussed by David Gauthier in "Assure and Threaten," *Ethics* 104 (1994):

What's more, the requisite justification is unlikely to emerge from an analysis of concepts such as "reason" or "rational action." As we have seen, these are formal concepts that have no application except in relation to a substantive object or standard of success. Because these concepts implicitly require such a standard to be supplied, we can hardly expect to deduce it from them, any more than we should expect to deduce the object of a game from the mere concept of winning, or the object of a hunt from the mere concept of a quarry.<sup>25</sup>

I cannot prove that the task of justifying an externalist conception of reasons is impossible, but I think it's going to be awfully hard. I'm just a fainthearted externalist, I guess. Being fainthearted, however, I want to consider whether the benefits of externalism can be obtained without the burdens. I shall therefore turn to an alternative conception of practical reason, which straddles the line between internalism and externalism.

#### OUTLINE OF AN ALTERNATIVE VIEW

Suppose that we want to frame a conception of reasons that isn't relativized to the inclinations of particular agents. That is, we want to identify particular things that count as reasons for acting simpliciter and not merely as reasons for some agents rather than others, depending on their inclinations.

One way to frame such a conception is to name some features that an action can have and to say that they count as reasons for someone whether or not he is inclined to care about them. The problem with the resulting conception, as we have seen, is that it entails the normative judgment that one ought to be inclined to care about the specified features, on pain of irrationality, and this normative judgment requires justification.

The advantage of internalism is that it avoids these normative commitments.<sup>26</sup> It says that things count as reasons for someone only

---

690–721. I discuss Gauthier's treatment of this notion in my "Deciding How to Decide," in Cullity and Gaut, eds.

25. Kant's conception of practical reason, as I understand it, is an attempt to circumvent this problem, by using the concept of a reason, not to identify which features are reasons, but rather to identify which features aren't, and by replacing the rule of acting for reasons with a rule of not acting for nonreasons. On Kant's conception, as I understand it, the object of practical reasoning is to act on any consideration but one whose being a reason would entail a contradiction. It's like a hunt whose object is to locate anything but that which could not possibly be a quarry. Perhaps the generic concepts of a quarry or a reason can indeed serve this modest role.

26. I don't mean to imply that internalism avoids all normative commitments. In "Skepticism about Practical Reason," Korsgaard suggests that the instrumental principle of adopting the means to one's ends is a substantive norm; she defends this point at length in "The Normativity of Instrumental Reason." But even if internalism has to justify requiring us to care about the means to what we already care about, it avoids

if he is inclined to care about them, and so it leaves the normative question of whether to care about them entirely open. Yet if we try to leave this question open, by defining things as reasons only for those inclined to care about them, we'll end up with a definition that's relativized to the inclinations of particular agents—won't we?

Not necessarily. For suppose that all reasons for acting are features of a single kind, whose influence depends on a single inclination. And suppose that the inclination on which the influence of reasons depends is, not an inclination that distinguishes some agents from others, but rather an inclination that distinguishes agents from non-agents. In that case, to say that these features count as reasons only for those who are inclined to care about them will be to say that they count as reasons only for agents—which will be to say no less than that they are reasons for acting, period, since applying only to agents is already part of the concept of reasons for acting. The restriction on the application of reasons will drop away from our definition, since it restricts their application, not to some proper subset of agents, but rather to the set of all agents, which is simply the universe of application for reasons to act.

\* \* \*

The foregoing paragraph is an outline for a conception of reasons for acting—a bare outline that needs filling in. The remainder of this article will be devoted to filling it in, at least to some extent, though not, I admit, to the extent that's needed. I shall begin by making a digression into the subject of theoretical reasoning. My hope is that we can understand reasons for acting by analogy with reasons for belief.<sup>27</sup>

The nature of reasons for belief, and the inclination that mediates their influence, are fairly clear. The object of theoretical reasoning is to arrive at true belief;<sup>28</sup> and since true belief needn't be defined in

---

the further and heavier burden of justifying any requirements to care about particular things.

27. The analogy between theoretical and practical reason is being pursued independently by my colleague Peter Railton, with somewhat different results. See his "What the Noncognitivist Helps Us to See the Naturalist Must Help Us to Explain," in *Reality, Representation, and Projection*, ed. John Haldane and Crispin Wright (New York: Oxford University Press, 1993), pp. 279–300, pp. 292 ff.; "A Kind of Nonsubjective Reason?" in *Essays in Honor of Kurt Baier*, ed. J. Schneewind (New York: Open Court, in press); and "On the Hypothetical and Non-hypothetical in Reasoning about Action," in Cullity and Gaut, eds. David Gauthier also discusses the analogy, but he ultimately rejects it ("Assure and Threaten," pp. 699–702).

28. Some may be inclined to think that the object of theoretical reasoning is not true belief but empirically adequate and explanatorily fruitful belief, or belief of some other kind. My argument doesn't depend on the outcome of this disagreement. What matters for my purposes is that theoretical reasoning aims at some outcome specified substantively (i.e., not in terms of its being the object of theoretical reasoning or belief).

terms of success in theoretical reasoning, it constitutes a substantive rather than formal standard of success.<sup>29</sup> Reasons for a particular belief are recognized by their perceived relevance to this substantive standard of success, as considerations that appear to guarantee or probabilify the truth of the belief.<sup>30</sup> And these considerations influence a person's beliefs by virtue of an inclination to believe what seems true. Here, then, are considerations of a single kind and a single inclination to mediate their influence.

Perhaps we should ask whether the absence of this inclination would undermine the existence of reasons for belief or would alternatively undermine the believer's claim to rationality.<sup>31</sup> The answer to this question would determine whether reasons for belief were internal or external reasons. If someone weren't inclined to believe what seemed true, would signs of truth in a proposition no longer count as reasons for him to believe it? Or would he no longer qualify as a rational believer?

\* \* \*

Both, I think—which goes to show that the question incorporates a false dichotomy. I shall argue that the dichotomy should be replaced with a subtler account of theoretical reasoning, along the following lines.

If someone isn't inclined to believe what seems true on a topic, he is no longer subject to reasons for believing things about it; but he is no longer subject to reasons for belief about it, I shall argue, because he is no longer a believer about it at all, and a fortiori no longer a rational believer.<sup>32</sup> He isn't in the business of forming beliefs on the

29. The claim that truth isn't defined in terms of success in theoretical reasoning is potentially controversial. It must be rejected by those who hold a pragmatist conception of truth as the eventual deliverance of rational inquiry. In my view, however, the pragmatist conception renders theoretical reasoning vacuous, like a game whose only object is winning.

30. In the case of inductive reasoning, of course, we may have trouble saying what relevance reasons have to the truth of a belief. Nevertheless, such reasons count as reasons for a belief because they make it seem true, even if we cannot say how or why. (An alternative way of handling this case would be to point out that inductive reasons satisfy a substantive procedural criterion of correctness in inductive inference. See n. 23 above.)

31. Williams raises this question and seems to suggest that the absence of an inclination toward the truth would undermine the existence of reasons for belief ("Internal Reasons and the Obscurity of Blame," p. 37). This is, of course, the internalist answer to the question.

32. I believe that Korsgaard makes a similar point ("The Normativity of Instrumental Reason," p. 42). In passages such as this, where Korsgaard seems to be pursuing a strategy like the one I am developing here, I begin to doubt whether she really is an externalist, in Williams's sense of the term. My reasons for this doubt will be explained in the text, below, when I explain why I do not regard the present strategy as a version of externalism.

topic, to begin with, unless he is inclined to believe what seems true about it.

Thus, reasons for believing something apply only to those who are inclined to believe what seems true on the topic, and so they are like internal reasons; but to say that they are reasons only for those who are so inclined is just to say that they are reasons only for potential believers on the topic—which is to say no less than that they are reasons for believing, period. Reasons for belief can therefore be identified independently of the inclinations of individuals, and so they are like external reasons, too.

The foregoing paragraph is a bare outline for an account of theoretical reasoning, and this outline also needs filling in. In order to fill it in, I shall have to explore the sense in which being inclined toward the truth is essential to being a subject of belief. I therefore turn to a different thesis associated with the name of Bernard Williams, the thesis that belief is an attitude that “aims at the truth.”<sup>33</sup>

#### THE CONSTITUTIVE AIM OF BELIEF

The grounds for this thesis emerge when we try to distinguish belief from the other propositional attitudes. One difference between belief and other attitudes is that it entails regarding its propositional object as true.

The difference between believing that *P* and desiring that *P*, for example, is that the former attitude treats *P* as a report of how things are, whereas the latter treats *P* as a mandate for how things are to become.<sup>34</sup> Desire takes its propositional object as representing *facienda*—things that aren’t the case but are to be brought about. By contrast, belief takes its propositional object as representing *facta*—things that are the case and in virtue of which the proposition is true.<sup>35</sup>

33. Bernard Williams, “Deciding to Believe,” in *Problems of the Self* (Cambridge: Cambridge University Press, 1973), 136–51.

34. This difference between belief and desire can be obscured by the fact that desiring that *P* entails desiring *P* to be true, just as believing that *P* entails believing it to be true. These locutions obscure the difference between belief and desire because they use the infinitive ‘to be’, which is required for indirect discourse, to replace what would be different moods of the copula in direct speech. In believing *P* to be true, one believes in its completed truth, as would be expressed by the indicative statement that *P* is true; whereas in desiring *P* to be true, one desires its to-be-completed truth, as would be expressed by the optative that it *be* true. Thus, although we can speak either of believing or of desiring *P* to be true, transposing these statements from *oratio obliqua* to *oratio recta* reveals an underlying difference in the relation that *P* is taken as bearing to the world.

35. The language used in this contrast should not be overinterpreted. To say that belief involves regarding a proposition as true, or that desire involves regarding it as to be made true, is simply to articulate our concepts of belief and desire as propositional attitudes. We express the fundamental similarity among these content-bearing mental

\* \* \*

This conception of belief is correct as far as it goes, but it doesn't go far enough. It's incomplete because regarding a proposition as true is involved in many cognitive attitudes, including not only belief but also other attitudes from which belief must still be distinguished. Assuming a proposition—say, for the sake of argument—entails regarding it as a report rather than a mandate, as a truth rather than something to be made true. Even imagining that *P* entails regarding it as a completed rather than a to-be-completed truth. One hasn't imagined that *P* unless one has regarded *P* as reflecting how things are, and hence as true. Yet to assume that *P* or imagine that *P* is not to believe it, and so regarding a proposition as true must not be sufficient for belief.<sup>36</sup>

Of course, there is a sense in which things that are merely assumed or imagined are not regarded as really true. But the relevant sense is not that they aren't regarded as true at all; it's rather that they are regarded as true but not really—regarded as true, that is, but not seriously or in earnest. What distinguishes a proposition's being believed from its being assumed or imagined is the spirit in which it is regarded as true, whether tentatively or hypothetically, as in the case of assumption; fancifully, as in the case of imagination; or seriously, as in the case of belief.

\* \* \*

What's the difference between seriously regarding a proposition as true and doing so in some other spirit? Here is the point at which belief is distinguished from other attitudes by its aim.

---

states by describing them as ways of regarding propositional contents, and we express the differences among them by differentiating among the ways in which those contents can be regarded. The resulting locutions should not be understood as positing any particular mental architecture, least of all an inner eye that squints at propositions or raises its eyebrow at them so as to regard them in different ways. Rather, these locutions simply translate our terms for propositional attitudes into a common vocabulary, in which their similarities and differences can be clearly expressed. To say that belief entails regarding a proposition as true is therefore not to commit ourselves to any particular theory about which physical, neurological, or otherwise subdoxastic states make up the mental state of belief. It commits us only to a view about what such states must amount to if they are to constitute belief—namely, that they must amount to the state of regarding a proposition as true. For recent discussions of this phenomenon, commonly called "direction of fit," see Lloyd Humberstone, "Direction of Fit," *Mind* 101 (1992): 59–83; and G. F. Schueler, "Pro-attitudes and Direction of Fit," *Mind* 100 (1991): 277–81. Note that I understand direction of fit somewhat differently from these and other authors. For a fuller treatment of the differences, see my "The Guise of the Good," *Nous* 26 (1992): 3–26; and n. 55 below.

36. For related discussions of the similarities and differences among these cognitive states, see Jennifer Church, "Judgment, Self-Consciousness, and Object Independence," *American Philosophical Quarterly* 27 (1990): 51–60; and Mark Leon, "Rationalising Belief," *Philosophical Papers* 21 (1992): 299–314.

The sense in which fantasies and assumptions aren't serious is that they entail regarding a proposition as true—or accepting the proposition, as I shall put it—without sensitivity to whether one is thereby accepting the truth. We assume a proposition when we regard it as true for the sake of thereby framing a possibility to be entertained in argument or inquiry and when we can therefore be said to accept it for polemical or heuristic purposes. We imagine a proposition when we regard it as true for the sake of thereby stimulating or vicariously satisfying our desires and when we can therefore be said to accept it for recreational or motivational purposes. But we believe a proposition when we regard it as true for the sake of thereby getting the truth right with respect to that proposition: to believe something is to accept it with the aim of doing so if and only if it really is true.

Thus, the purpose or aim with which a proposition is regarded as true is partly constitutive of the resulting attitude toward the proposition. It determines whether the proposition is being accepted hypothetically, as in assumption; playfully, as in imagination; or seriously, as in belief. These attitudes can therefore be conceived as having two tiers. The first tier, which they share and by virtue of which they differ as a group from the conative attitudes, is the attitude of regarding a proposition as true—the attitude of bare acceptance. The second tier, in which the various cognitive attitudes differ among themselves, encompasses the different aims with which a proposition can be accepted.<sup>37</sup>

To say that our attitude toward a proposition is partly constituted by the aim or purpose with which we accept the proposition is not to say that the aim is itself an attitude of ours, or that acceptance is an action. This point cannot be overemphasized.<sup>38</sup> Acceptance is a mental

37. An example that can help to illustrate this conception of the propositional attitudes appears in Bernard Williams's discussion of "Imagination and the Self" (in *Problems of the Self*, pp. 29–31). Williams compares two men who imagine assassinating the Prime Minister in the person of Lord Salisbury. One man imagines assassinating the Prime Minister but falsely believes that Lord Salisbury occupies that position; the other man, who knows that Lord Salisbury isn't Prime Minister, nevertheless imagines him to be, while also imagining a similar assassination. "On the purely psychological level," Williams remarks, "the same visualisings, the same images, could surely occur in both cases. The difference lies rather in how the story is meant" (p. 31). According to my account, "how the story is meant" should be understood in terms of the aim with which it is regarded as true that Lord Salisbury is Prime Minister. Each subject includes this identification in his "story," and thereby regards it as true. But one subject regards it as true for the sake of correctly identifying the Prime Minister, whereas the other regards it as true for the sake of his own entertainment.

38. The point will be lost on those who believe that any goal-directed movement, mental or physical, automatically qualifies as an intentional action. I reject this view, as will become clear on pp. 715 ff. My reasons for rejecting it are developed more fully in my "Guise of the Good" and "What Happens When Someone Acts?" *Mind* 101 (1992): 461–81. In any case, the present account of belief will be misunderstood if aims are assumed to be necessarily agential.

state whose aim may be emergent in the cognitive mechanisms by which that state is induced, sustained, and revised. For example, if our acceptance of a proposition is regulated by mechanisms performing their function of therein framing a possibility to be tested, then our acceptance may have a heuristic aim whether or not we have heuristic motives or take any action toward heuristic ends. Similarly, if our acceptance of a proposition is regulated by mechanisms performing their function of therein tracking the truth of the proposition, then it may have an epistemic aim whether or not we have or act on such an aim.<sup>39</sup> In short, our acceptance of a proposition may be aimed at the truth by our cognitive faculties rather than ourselves.

\* \* \*

This possibility suggests that one can have beliefs—aimed, as required, at the truth—while also being indifferent, at another level, to the truth of those beliefs. There are two ways of being indifferent to the truth, of which only one is an obstacle to believing.

To begin with, I can accept a proposition in a manner indifferent to its truth, thereby forming an assumption or fantasy rather than a belief. I am not then proceeding with indifference to the truth of a belief; I'm proceeding with indifference to the truth of what I accept, thereby falling short of belief altogether.

In another sense, however, I can be indifferent to the truth of something conceived as a belief. I cannot believe something without accepting it seriously—in an attempt, by me or my cognitive faculties, to arrive at acceptance of the truth—but I can still have further, second-order goals with respect to this attempt. For example, I can try to ensure that an attempt to accept what's true with respect to a proposition will lead to acceptance of that proposition whether it's true or not. This second-order attempt, to manipulate the outcome of a first-order attempt to accept what's true, is precisely what I undertake when I try to get myself to hold a particular belief irrespective of its truth. And in this case I am indifferent to the truth specifically of a

39. As David Phillips has pointed out to me, the mechanisms whose function is to track the truth may employ assumptions or even fantasies along the way. Thus, whether a particular instance of acceptance is an hypothesis, fantasy, or belief cannot depend on the ultimate aim toward which it is directed. Rather, the nature of each acceptance must depend on its immediate aim, as I have tried to indicate with the words 'therein' and 'thereby': to assume that *P* is to accept *P* for the sake of thereby formulating a possibility to be tested, whereas to believe that *P* is to accept *P* for the sake of thereby accepting the truth with respect to *P*. (Peter Railton raises the same problem in his "Truth, Reason, and the Regulation of Belief," *Philosophical Issues* 5 [1994]: 71–93.)

belief, because my indifference is directed at the success of something conceived as an attempt at accepting the truth.<sup>40</sup>

\* \* \*

I can thus fail to care about the truth of my beliefs. Yet if indicators of truth in a proposition are reasons for believing it, then indifference to the truth of my beliefs would seem to leave me insensitive to reasons, and hence irrational. My conception of theoretical reason would thus seem to resemble externalist conceptions of practical reason, in mandating a particular concern or inclination as required for rationality.

But this appearance is misleading. The conception of reasons for belief as indicators of truth doesn't imply that indifference to the truth of my beliefs would be irrational. Indifference to the truth of my beliefs would not in fact make me insensitive to the associated reasons for believing.

Of course, evidence for the truth of some belief may not sway me toward wanting or getting myself to hold that belief, if I'm indifferent to its truth. But we don't necessarily think that indicators of truth are reasons for such second-order measures as wanting or getting myself to hold beliefs. We identify them as reasons for *believing*, which are simply reasons for accepting something in the course of an attempt to arrive at acceptance of what's true. And insofar as I or my cognitive faculties attempt to arrive at the truth on a topic, that attempt will already make me potentially sensitive to indicators of the truth; whereas if no such attempt is in the works, the topic will be one on which I am not in the business of holding beliefs, in the first place.

What provides my sensitivity to reasons for believing, then, is not a second-order aim of having true beliefs but rather the first-order aim that makes my acceptance of something into a belief. And if this first-order aim is lacking from my approach to some topic, then I am not irrationally insensitive to reasons for belief about it; I am out of the business of having beliefs about it altogether, and so I am no longer subject to reasons for belief about it at all. Thus, my conception of theoretical reason doesn't condemn this form of indifference as irrational, either.

40. As Williams noted in "Deciding to Believe," this account of indifference to the truth of a belief explains the difficulty of acting on that indifference. In order to end up believing the proposition that I want to believe, I must accept it in the course of an attempt to accept what is true, not an attempt merely to accept this proposition. Indifference to the truth must not seep into my first-order attempt from my second-order attitude toward its success or failure. Some psychological partitioning is therefore necessary. On the difficulty of manipulating beliefs, see also Leon.

\* \* \*

In identifying something as a reason to believe a proposition, we are implicitly identifying it as a reason for a potential believer, someone who is in a position to believe or disbelieve the proposition at issue. Now, someone can be in a position to form a belief even though he lacks an interest in the truth of that belief—the second-order interest in the success of this attempt at accepting what's true. But he is not prepared to believe or disbelieve a proposition if he isn't prepared for an attempt to accept what's true with respect to it. Thus, he is not a potential believer with respect to a proposition—and hence not subject to reasons for believing it—in the absence of an inclination that would cause him to be swayed by indicators of its truth.

So when we say that indicators of truth are reasons for belief, we aren't making a normative judgment about whether to be inclined toward the truth; we're saying that they're reasons for someone only if he is inclined toward the truth, since we're identifying them as reasons of a kind whose universe of application is the set of potential believers, who are constitutively truth inclined. The question whether to be inclined toward the truth on some topic—and hence whether to be subject to reasons for belief about it—is left entirely open.

\* \* \*

In some sense, theoretical reasoning now seems to fit the model of internalism. Indicators of truth count as reasons for someone to believe only if he has a cognitive inclination that makes him susceptible to their influence. And reasons that apply to someone only if he's susceptible to their influence are supposed to be internal reasons.

At this point, however, the distinction between internal and external reasons is out of its depth, so to speak. Reasons for belief are dependent on a particular inclination, all right, but they're dependent on that inclination which makes one a believer. They don't depend on one's peculiar inclinations *as* a believer—on one's second-order attitudes toward or preferences among beliefs.

Indeed, the dependence of theoretical reasons on a cognitive inclination does not justify relativizing them to particular believers at all. The inclination on which these reasons depend is constitutive of belief itself, and to that extent they are reasons simply *for belief* rather than for any particular person to believe.<sup>41</sup> If something counts in a particular epistemic context as a reason to believe that *P*, then it counts in that

41. Of course, reasons for belief are also relative to an informational context, and insofar as different people are in possession of different information, they will be subject to different reasons. But if the informational context is held constant, the relativity of reasons to persons disappears.

context as a reason simply to believe that *P*, and not just for this or that believer to do so, since all potential believers of *P* are alike in the cognitive inclination that gives application to such a reason.

The question of whether reasons for belief are internal or external reasons thus presents a false dichotomy. Reasons for belief are like internal reasons in that they exist and exert an influence only in relation to a particular inclination; but they are like external reasons in that the inclination on which they depend is embedded in the attitude of belief, so that they can count as reasons for belief *per se*, in abstraction from motivational differences among individual believers.

\* \* \*

Maybe the way to understand the status of reasons for belief is to consider an analogy between belief and another enterprise that's partly constituted by a substantive aim. Consider reasons for sacrificing a pawn in the game of chess.

Reasons for sacrificing a pawn depend for their existence on a goal or aim, and in this respect they look like internal reasons. But the goal on which reasons for this move depend is partly constitutive of the move itself, because sacrificing a pawn is by definition a move in the game of chess, which is partly defined by its object; and because the move is by definition a sacrifice, which it can be only in relation to the object of the game. Reasons for sacrificing a pawn therefore exist in abstraction from the temperament of any particular player: they are reasons simply for the move itself, and in this respect they look like external reasons, too.

A player may have second-order aims with respect to his success or failure in a particular game of chess. He may even have the goal of losing a game—if his opponent is a sensitive eight-year-old, for example. But in order to lose a game of chess, he must stay in the game, by continuing to pursue its object, however insincerely or ineffectually. And so long as he is pursuing that object, he will have the inclination that answers to reasons for sacrificing a pawn.

A player can lose his susceptibility to those reasons only by giving up the associated object—moving his pieces around aimlessly, for example. In that case, he will in effect have quit the game: his opponent will say, not just “You’re letting me win,” but “You’re not playing any more.” Once the player has quit the game of chess, however, he has quit the only game in which pawns can be sacrificed, and his resulting insensitivity to reasons for sacrificing a pawn will not make him irrational. To someone who isn’t playing chess, reasons for sacrificing a pawn simply don’t apply.

In sum, reasons for sacrificing a pawn apply to anyone with the capacity to do so, irrespective of his inclinations about how to exercise

that capacity. They apply to him only because he has an inclination that lends them an influence, of course, but the requisite inclination is the one that makes him a chess player, not one that determines his individual style of play.

#### APPLYING THE ANALOGY TO PRACTICAL REASONING

I think that practical reasoning occupies the same middle ground between internalism and externalism. That is, reasons for acting apply to someone only because he has an inclination that lends them an influence, but the requisite inclination is the one that makes him an agent, not one that determines his individual course of action.

This account of practical reason simply follows the structure of theoretical reason, as analyzed above. That analysis began with the claim that belief is distinguished from other cognitive states by a substantive goal, and then it claimed that an inclination toward this goal creates the susceptibility necessary to the application of reasons for believing. Perhaps, then, action can be distinguished from other forms of behavior by a substantive goal, and an inclination toward this goal can create the susceptibility necessary to the application of reasons for acting. In that case, reasons for acting would be considerations relevant to the constitutive aim of action, just as reasons for believing are indicators of truth, which is the constitutive aim of belief. And anyone who wasn't susceptible to reasons for acting, because he had no inclination toward the relevant aim, wouldn't be in a position to act, anyway, and therefore wouldn't be subject to reasons for acting; just as anyone who has no inclination toward the truth isn't in a position to believe and isn't subject to reasons for belief.

The account rests, of course, on the initial claim that behaviors qualify as actions by virtue of having a particular aim. Let me say a word about the philosophical point of such a claim.

The point of specifying which behaviors qualify as actions is not, I think, to delineate the extension of 'action' or 'to act' as used in ordinary language. These terms are used quite loosely, in application not only to paradigm cases of action, in which human agency is exercised to its fullest, but also to marginal cases, in which agency is exercised only partially or imperfectly. The fundamental question in the philosophy of action is not how imperfect an exercise of agency can be while still qualifying as an action. The question is the nature of agency itself, and agency, like any capacity, fully reveals its nature only when fully exercised. We therefore want to know what makes for a paradigm case of action, a full-blooded action, an action par excellence.

I claim that what makes for an action, in this sense, is a constitutive aim. This claim sounds odd, to say the least. We may think that a full-

blooded action must have some goal or other;<sup>42</sup> but we tend to think that its status as an action doesn't depend on what goal it has. Action, we tend to think, is just behavior aimed at some goal, any goal.

\* \* \*

In my opinion, however, we are mistaken in assuming that behavior approaches full-blooded action by having a goal of the sort that varies from one action to another. Simply being goal directed is not a mark of action.

Consider a case of unintentional behavior. An old friend unexpectedly walks into your office, and surprise lights up your face: your eyes widen, a smile flashes, an exclamation escapes your lips. These reactions just happen to you, and they may even hit you with an aftershock of surprise. Now suppose, instead, that you encounter your friend on the quad, recognizing him as he approaches. You are moved to the same reactions, but you now have a chance to modulate them or compose them into an intentional expression of surprise.

Take another case of unintentional behavior. Say, a child accidentally brushes a glass off of the table, and your hand shoots out to catch it. Everything happens so fast that you see your hand catching the glass before you fully realize that the glass is falling. Now suppose, finally, that another child—an older and sassier child—hefts the glass with a smirk and calls, "Here, catch!" You then undertake the same behavior, but as a fully intentional action.

The first instance in which you catch the glass is an instance of behavior directed at a goal, but it isn't a full-blooded exercise of your agency. Unlike your reflexive expression of surprise, which springs *out of* the emotion of surprise but not *toward* any purpose or goal, the reflexive extension of your hand is aimed at something—namely, preventing the glass from smashing on the floor.<sup>43</sup> Despite being goal directed, however, this behavior still lacks some element that's necessary to full-blooded action. So what makes for action is not simply being goal directed.

\* \* \*

The question is what's missing from this goal-directed behavior. In my view, what's missing is some additional goal that every action shares, no matter what its other, contingent goals may be.

42. See, e.g., Jay Wallace, "How to Argue about Practical Reason," *Mind* 99 (1990): 355–85, p. 359: "To act intentionally . . . is necessarily to be in a goal-directed state"; see also Smith, "The Humean Theory of Motivation."

43. The idea that some actions spring out of motives without being directed toward any ends or goals is defended by Michael Stocker, "Values and Purposes: The Limits of Teleology and the Ends of Friendship," *Journal of Philosophy* 78 (1981): 747–65.

There is an ancient thesis along these lines, to the effect that action, no matter where it aims, must thereby aim at the good.<sup>44</sup> This thesis identifies a constitutive goal of action—the good—and it thus implements the strategy of analysis that I favor. But in this implementation, the strategy fails to achieve its purpose, since it doesn't avoid the twin pitfalls of internalism and externalism.<sup>45</sup>

The thesis that action constitutively aims at the good can be interpreted in at least two ways. It may simply mean that an action must aim at something, which consequently counts as good in the sense of being that whose attainment will make the action a success. But this sense of the word 'good' is a formal sense, denoting whatever is the aim of an action. It identifies no particular thing at which every action must aim, and hence no particular kind of consideration as capable of influencing anyone insofar as he is an agent. If the thesis uses this formal sense of the word 'good', then the considerations that it classifies as reasons will vary along with the good being aimed at. Reasons will then depend for their application on one's inclinations as an agent, as they do under internalism.

In order to avoid this consequence of internalism, the ancient thesis would have to identify a substantive goal for action, by saying that every action aims at something conceived as good in a sense independent of its being the aim. But when the thesis uses a substantive sense of the word 'good' in this manner, it characterizes action as necessarily well-intentioned, thus ruling out various kinds of perversity. To those who believe, as I do, that behavior can still qualify as action even if its end-in-view is conceived as bad, the thesis will now appear to be burdened with controversial normative commitments, like the version of externalism considered above.<sup>46</sup>

\* \* \*

The ancient thesis goes wrong, I think, in treating the constitutive aim of action as something shared or jointly promoted by all of an agent's other ends-in-view, as if it were an ultimate or all-encompassing end. If action is to be constituted by an aim, however, that aim cannot be an end at all.

44. This view is echoed by Donald Davidson in "How Is Weakness of the Will Possible?" in *Essays on Actions and Events*, pp. 21–42, p. 22.

45. This problem was foreshadowed in the text accompanying n. 22, above.

46. In adopting an evil end, the perverse agent may of course be said to make evil his good, as Satan does in *Paradise Lost* (bk. 4, line 110). But Satan makes evil his good only in the formal sense that its attainment will be the criterion of his success. The fact that even Satan's actions aim at the good in this formal sense doesn't help us to identify a substantive aim that constitutes them as actions. G. E. M. Anscombe discusses this passage (*Intention* [Ithaca, N.Y.: Cornell University Press, 1963], p. 75), and I have elsewhere criticized her discussion ("Guise of the Good," pp. 18 ff.).

An end is something conceived by an agent as a potential object of his actions. It is therefore something that one cannot have unless one already is an agent, in a position to act, and so it cannot be something that one must already have in order to occupy that position. If action is to be constituted by an aim, that aim must be, so to speak, subactional or subagental—something that a subject of mere behavior can have, and by having which he can become an agent, as his behavior becomes an action.

This subactional aim can be discerned, I think, in our contrasting pairs of behaviors. It is that which the unintentional behaviors are missing in comparison with the corresponding full-blooded actions.

\* \* \*

Intuitively speaking, what these behaviors lack is that, while directed *at* various things, they are not directed *by* you. When the glass is brushed off the table, for example, behavior aimed at arresting its fall is initiated and completed before you know it, and so you have no chance to take control of that behavior. In the intentional instance, the same goal-directed behavior occurs, but it occurs under your control.

The kind of control at issue here is not the sensorimotor process that adjusts ongoing behavior in light of perceived progress toward a desired outcome. That process of real-time adjustment is simply eye-hand coordination, which occurs in both the intentional and the reflexive cases. What's missing from the reflexive case is conscious direction on your part, which is something other than eye-hand coordination. When goal-directed behavior proceeds under this conscious control, it becomes a full-blooded action, rather than a well-coordinated reflex.<sup>47</sup> And behavior that isn't directed at a goal can become an action in the same fashion. The smile that springs spontaneously from your emotion of surprise isn't aimed at any result, but it, too, can be transformed into a full-blooded action if it is brought under your conscious control.

Now, if an action comprises behavior of which you take control, then taking control of your behavior cannot itself be an action; otherwise, a vicious regress will ensue. Yet controlling your behavior is indeed an activity: it's something that you do. The reason why the falling glass leaves you no time to perform a full-blooded action is that, although it leaves you time to stick out your hand, it doesn't leave you time to do something else that's essential to a full-blooded

47. I do not mean that every part or aspect of the behavior must come under your conscious control in order for the behavior to constitute a full-blooded action. How you execute the catch may still be left to those reflexes which make up your skill as a catcher; that you execute a catch, however, must come under your control, or the catch won't be an action in the fullest sense.

action—that is, to exercise conscious control of your catch. Hence when you catch the glass intentionally, you must be doing two things: extending your hand in order to avert a mishap and exercising control over that behavior.

Let me reiterate that I am using the noun ‘activity’ and the verb ‘to do’ in senses that do not imply the performance of a full-blooded action. To suggest that an action comprises behavior on which you perform the action of exercising control would be absurd. But you do many things that aren’t actions—such as when you reflexively stick out your hand to catch a falling glass or smile out of surprise. And exercising conscious control over your behavior is indeed something that you do, in this thin sense of the verb.

I therefore suggest that our ordinary concept of a full-blooded action is in fact the concept of two, hierarchically related activities. Action is like the corporate enterprise of work performed under management: it’s behavior executed under conscious control. And just as the corporate enterprise includes both a basic work activity and the higher-order activity of managing that work (neither of which is itself a corporate enterprise), so full-blooded action comprises both a basic activity and the higher-order activity of controlling it (neither of which is itself an action).

\* \* \*

This analysis of action suggests how action might have a constitutive goal. According to the analysis, various actions involve various behaviors—directed, in many cases, at various goals—but they also share an additional, higher-order activity, the activity of consciously directing these behaviors. This activity is constitutive of action, in the sense that its addition is what makes a full-blooded action out of a merely reflexive or unintentional movement. If this higher-order activity entails the pursuit of a goal, then there may indeed be a constitutive goal of action.<sup>48</sup>

What I have in mind here is not an ulterior goal or aim toward which behavior is consciously directed, as a corporation’s work activity might be managed toward the end of maximizing profits. The executive officers can still manage the work of a corporation without having the goal of profit maximization, in particular, and so this goal is not itself essential to work’s being performed under management. Similarly, a goal toward which behavior is consciously directed may not be

48. I have elsewhere presented an independent argument for this thesis (“What Happens When Someone Acts?”). The idea that practical reason has motives of its own, directed at the control of one’s behavior, is contained in the theory of motivation attributed to Plato by John Cooper, “Plato’s Theory of Human Motivation,” *History of Philosophy Quarterly* 1 (1984): 3–21.

essential to the behavior's being consciously directed, insofar as behavior might be consciously directed at other goals or no goal in particular.

What I have in mind is a goal that must be pursued if behavior is to be consciously directed at all. This goal will not be one of the agent's ends-in-view, nor will it be something on which those ends converge. Rather, it will be something whose pursuit is ancillary to theirs—something whose pursuit transforms them, from outcomes sought unconsciously or reflexively, into ends at which action is consciously directed.

### THE CONSTITUTIVE GOAL OF ACTION

What is this goal? A hint lies in the fact that consciously controlling one's behavior is not something that one can do without aiming to.<sup>49</sup> Maybe, then, the aim without which there is no conscious control of behavior is simply the aim of being in conscious control of one's behavior. If so, then the constitutive aim of action will turn out, in Kantian fashion, to be autonomy.<sup>50</sup> And considerations will turn out to qualify as reasons—also in Kantian fashion—by virtue of their relevance to our autonomy rather than their relevance to our interests or our good.<sup>51</sup>

These remarks are merely suggestive at best, and this is not the place to develop them into a full account of autonomy or its role as the constitutive goal of action.<sup>52</sup> I can only sketch how they might be developed.

\* \* \*

My sketch begins with the conception of autonomy as conscious control over one's behavior. Consciously controlling one's behavior in-

49. Bernard Williams has pointed out to me that one can consciously control one's behavior while aiming not to—as, for example, when one unsuccessfully tries to let one's reflexes or instincts take over. But this point strikes me as compatible with my claim that one cannot consciously control one's behavior without aiming to. Trying not to control one's behavior involves a second-order goal, of relaxing one's first-order efforts at control. If one continues to control one's behavior while trying not to, the reason is that one continues to aim at controlling it while trying not to persist in that aim. (Remember that the aims under discussion here may be subagential. See pp. 716–17, above).

50. Thanks to Chris Korsgaard for publicly daring me to express this thought.

51. Stephen Darwall has proposed a similar conception of reasons, under the name 'autonomist internalism' ("Autonomist Internalism and the Justification of Morals," *Nous* 24 [1990]: 257–68). Of course, considerations may be relevant to our autonomy because of their relevance to our interests. The point is that their relevance to autonomy will be what makes them reasons for acting.

52. See my *Practical Reflection* (Princeton, N.J.: Princeton University Press, 1989); and "The Story of Rational Action," *Philosophical Topics* 21 (1993): 229–54.

volves two elements: being conscious of one's behavior and controlling it. How are these elements connected?

One possibility is that they aren't connected at all. Conscious control might just be the sum of two independent elements, control over what one is doing and consciousness of what one is doing. Another possibility is that exercising control over one's behavior is what brings it to consciousness. One might control what one is doing and thereby become conscious of that behavior.

The problem with these possibilities is that they would leave an agent's knowledge of his behavior dependent on the usual inbound channels, such as perception of the behavior itself or introspection on the process by which it is directed. And as many philosophers have noted, an agent's knowledge of his behavior is not receptive knowledge: an agent knows what he is doing, as they say, without observation.<sup>53</sup>

The work of these philosophers points to a third possibility for the relation between self-control and self-awareness. Maybe consciousness of what one is doing is that by which one exerts control. Consciously controlling one's behavior would then be—not just controlling it and also, or thereby, becoming conscious of it—but rather having a *controlling consciousness* of one's behavior, a guiding awareness of what one is doing. This possibility would account not only for an agent's self-control but also for the quality of his self-awareness, since his knowledge of what he was doing would be, so to speak, directive rather than receptive knowledge.<sup>54</sup>

\* \* \*

But how can knowledge be directive? For the answer, let me return to my earlier account of cognition. (I'll give the answer in this section and then illustrate it in the next.)

Consciousness or knowledge must be a cognitive state, and so it must involve regarding propositions as true rather than as to be made

53. See Anscombe, *Intention*; Stuart Hampshire, *Freedom of the Individual* (Princeton, N.J.: Princeton University Press, 1975), chap. 3; Brian O'Shaughnessy, *The Will: A Dual Aspect Theory* (Cambridge: Cambridge University Press, 1980), chap. 8. See also Ludwig Wittgenstein, *Philosophical Investigations*, trans. G. E. M. Anscombe (Oxford: Blackwell, 1967), secs. 627 ff.

54. Compare the ancient and medieval notion of "practical knowledge," which is "the cause of what it understands" (Aquinas, *Summa Theologica*, Ia IIae, Q3, art. 5, obj. 1). Anscombe discusses this notion in the last two paragraphs of her paper "Thought and Action in Aristotle" (in *New Essays on Plato and Aristotle*, ed. by R. Bambrough [London: Routledge & Kegan Paul, 1965], 143–58), thereby picking up a theme that was left undeveloped in *Intention*, pp. 1–5, 56–58, 87. See also David Pears, *Motivated Irrationality* (Oxford: Clarendon, 1984), chap. 8; and Arthur Danto, "Action, Knowledge, and Representation," in *Action Theory*, ed. by Myles Brand and Douglas Walton (Dordrecht: Reidel, 1976), pp. 11–25.

true. It must also be a serious cognitive state, regarding propositions as true in an attempt thereby to get the truth right. Indeed, the success that's implied in the concept of consciousness or knowledge is success in this very attempt, to regard as true what really is true.

But there are two ways of attempting to regard as true, or accept, what really is true. One way is to accept a proposition in response to its being true; the other is to accept a proposition in such a way as to make it true. Note that the latter method does not entail regarding the proposition as *to be made* true. It entails attempting to make the proposition true by regarding it as such, but attempting to make a proposition true by regarding it as true is quite different from regarding it as to be made true. The proposition is regarded as fact, not *faciendum*, and so it is accepted, in a cognitive rather than conative attitude. What's more, the proposition is accepted seriously, not hypothetically or frivolously. For in attempting to accept something so as to make it true, one attempts to reach the position of accepting a genuine truth, no less than when one attempts to accept something in response to its being true. In either case, one's acceptance aims at correspondence between what's regarded as true and what is true, and so it is a serious cognitive attitude, whose success deserves to be called knowledge.<sup>55</sup>

How can one regard a proposition as true in such a way as to make it true? Well, when one accepts a proposition in response to its truth, one registers the influence of evidence and other reasons for belief, thereby manifesting an inclination to conform one's acceptance to the facts. Accepting a proposition in such a way as to make it true would simply require a converse inclination, to conform the facts to

55. Here I am expanding on two themes that I have discussed elsewhere. First, I am expanding an earlier critique of the traditional notion of direction of fit (in my "Guise of the Good"). In my view, this notion conflates two different distinctions. One is the distinction between the cognitive and the conative—the distinction between accepting, or regarding as true and approving, or regarding as to be made true. The other is a distinction between the receptive and the directive, which are two different ways of attempting to accept what's true—namely, by accepting so as to reflect the truth, and by accepting so as to create the truth. If these distinctions are conflated under the heading 'direction of fit', then one and the same mental state can appear to have two different directions of fit, since a subject can attempt to accept what's true by accepting something so as to make it true. The resulting state is cognitive rather than conative, but directive rather than receptive: it's directive cognition. I would claim that this state of directive cognition is the state of intending to act. This is the second theme on which I am currently expanding. In the past, I have said that an intention is a self-fulfilling and self-referring belief (*Practical Reflection*, chap. 4; see also "How to Share an Intention," *Philosophy and Phenomenological Research* 57 [1997]). The present discussion explains why I call it a belief, but also why I can dispense with that label. What matters is that intention is a state of directive cognition, not whether that state should be called belief.

one's acceptance. And one can indeed be inclined to conform the facts to one's acceptance, if the proposition accepted is about one's own behavior. One need only be inclined to do what one accepts that one will do. If one has this inclination, then accepting that one will do something can be a way of making this proposition true, and it can therefore be an attempt at accepting the truth.

\* \* \*

This admittedly convoluted proposal can be applied to the contrast between your reflexively and intentionally catching that glass. In both cases, your desire to save the glass causes your hand to extend. In both cases, you're aware of this causal sequence, since you're aware of extending your hand in order to save the glass. But in only one of the cases is your knowledge directive, or your behavior autonomous.

When you extend your hand reflexively, you react before you know it, but then you observe your reaction. Extending your hand in order to save the glass causes you to accept the proposition that you're doing so. When you react intentionally, however, acceptance precedes behavior: you accept that you'll extend your hand to save the glass, and this acceptance is what prompts you to do so.

In the latter case as well as the former, your acceptance is an attempt to accept something true. You're not just hypothesizing or fantasizing that you'll extend your hand: you're seriously regarding it as true that you will extend it. Of course, your acceptance of this proposition is not an attempt to accept something that's true antecedently; it's an attempt to accept something whose truth will follow as a result. But it is not therefore less serious as an attempt to accept a truth. If the proposition accepted comes true, then its acceptance is a cognitive success—an instance of directive knowledge.

In sum, instead of reacting before you know it, you react after and because you know it, and that's what makes your behavior an autonomous action. You act autonomously because you extend your hand in, and out of, a knowledge of what you're doing.

\* \* \*

But why would your extending a hand to save the glass result from your accepting that you would do so?

Suppose that you have an inclination toward being in conscious control of your next move. This inclination will inhibit you from doing anything out of other motives until you've accepted that you're going to—precisely so that you'll do it only after and because you know it, and hence under conscious control. Once you accept that you're going to do something, however, the inclination toward being in conscious

control will reinforce your other motives for doing it, since doing what you've accepted you'll do is what puts consciousness in control. Your inclination toward conscious control is thus converted, from an inhibition against doing something into a motive in favor of doing it, by your accepting that you'll do it. Accepting that you'll extend your hand to save the glass can therefore prompt you to do so.

Here, then, is how autonomy can serve as the constitutive goal of action. The goal-directed movement of your hand comes under your conscious control because it is prompted by your accepting that you will perform such a movement. And it is prompted by that acceptance because of your inclination toward conscious control of what you're doing—which is just an inclination toward autonomy. Your movement thus becomes autonomous precisely by manifesting your inclination toward autonomy; and in becoming autonomous, it becomes a full-blooded action. A full-blooded action is therefore behavior that manifests your inclination toward autonomy, just as a belief is a cognitive attitude that manifests your inclination toward the truth.

\* \* \*

My view is that your inclination toward the constitutive goal of action also mediates the influence of your reasons for acting, just as your inclination toward the truth mediates the influence of your reasons for belief. Your reasons for acting can be displayed as the premises of a practical inference:

I want to save that glass.  
I could save the glass by extending my hand.  
So I'll extend my hand.

Since the premises of this inference are about how to fulfill a desire of yours, they sound like reasons that Williams would call internal. But in my view, they don't influence you in quite the way that internal reasons are supposed to.

Here is how internal reasons are supposed to work. The first premise of your inference is about a desire of yours: "I want to save that glass." The second premise is about the means to the object of that desire: "I could save the glass by extending my hand." The desire mentioned in the first premise and the belief expressed in the second combine to motivate the action mentioned in your conclusion: "So I'll extend my hand." According to the internalist tradition, this process of motivation is the very process whereby reasons for acting exert their influence as reasons.

This conception of how reasons exert their influence encourages a particular reading of the statements displayed above. Since the influence of reasons is conceived as the motivational influence of desire and belief,

and since the second premise expresses the operative belief, the first premise is read as expressing the associated desire.<sup>56</sup> Similarly, the conclusion is sometimes read as expressing—or standing in for—the action itself, which is said to be the real conclusion of your inference.<sup>57</sup> The three displayed statements are thus interpreted as expressions of your reasons and of the action that they influence you to perform.

I don't dispute the traditional account of how desire and belief motivate behavior. My quarrel is with the claim that when desire and belief motivate behavior, they exert the influence of reasons.<sup>58</sup> You extend your hand, I agree, out of a desire for something and a belief about how to attain it. But you can extend your hand out of a desire and belief even when you do so reflexively, without knowing what you're doing or why, and hence without the benefit of practical reasoning.

In my view, extending your hand out of a desire and belief is the underlying behavior over which you may or may not exercise conscious control—the underlying work that may or may not come under your executive management. And practical reasoning is the process by which you exercise conscious control over this activity in some cases but not others. If you extend your hand without any guiding knowledge of what you're doing, then even though your behavior is motivated by a desire and belief, it isn't under your conscious control, and so it isn't a full-blooded action. Your behavior amounts to a full-blooded action only when it is performed in, and out of, a knowledge of what you're doing—or, as I have said, after and because you know it.<sup>59</sup>

\* \* \*

This view of practical reasoning encourages a different interpretation of the statements displayed above. The first premise expresses a desire-

56. See Davidson, "Intending," in *Essays on Actions and Events*, pp. 83–102, p. 86. Because Davidson thinks that this premise should express your desire, he would reformulate it, from "I want to save that glass" to "Saving that glass would be desirable."

57. Davidson, "How Is Weakness of the Will Possible?" p. 32; and "Intending," pp. 98–99.

58. This quarrel is a continuation of my "Guise of the Good" and "What Happens When Someone Acts?"

59. Note that in my account, your autonomy isn't an ability to control the motions of your hand; it's an ability to control your behavior, which is bodily motion psychologically understood, in terms of its motivation. Even a robot can control whether its hand moves. It takes an autonomous agent to control whether he moves his hand out of a desire to save a glass. The object of autonomous control is thus the entire behavior, comprising motivation as well as movement. The same point can be put in (somewhat) Kantian terms, as follows. Acting autonomously isn't just moving in accordance with one's idea of a movement; it's acting in accordance with one's idea of a law—in this case, the law of motivation.

based reason, in my view, but the reason expressed is not the desire itself. The reason expressed by “I want to save that glass” is your recognition of the desire.

This recognition is a reason because, together with the belief expressed in your second premise, it forms a potentially guiding awareness of what you would be doing in extending your hand. The awareness that you want to save the glass, and that extending your hand would save it, puts you in a position to frame a piece of directive knowledge—“I’m extending my hand in order to save the glass”—a proposition that you can now make true by accepting it. Your awareness of the desire thus presents the behavior of extending your hand in a form prepared for your conscious control, as a potential object of your directive grasp. It presents the behavior, if you will, as fit for (en)action, given the constitutive aim of action, just as theoretical reasons present a proposition as fit for belief, given the constitutive aim of belief.

This view of practical reasoning requires far more elaboration and defense than I can offer here. Its only relevance to this article is that it implements the compromise that I favor between internalism and externalism. For according to this view, even desire-based reasons for acting derive their influence from an inclination other than the desires on which they are based.

\* \* \*

The reasons displayed above are desire based in the sense that they mention your desire to save the glass and the means to fulfilling that desire. Yet their influence as reasons is not mediated by the desire that they mention.

Your desire to save the glass does exert a motivational influence in this example. But its influence as a motive contributes to the underlying activity of extending your hand in order to save the glass—the activity that comes under the control of your practical reasoning. And the contribution of your reasons to the control of this activity is distinct from the contribution of your motives to the activity itself.

What exerts the influence of a reason in this example is the recognition that you want to save the glass. And this recognition doesn’t influence you by engaging your desire to save the glass. Wanting to save the glass is a motive that can be engaged by considerations about how to save it, not by the recognition that you want to. The recognition that you want to save the glass engages a different inclination, your inclination toward autonomy—toward behaving in, and out of, a knowledge of what you’re doing. And it thereby exerts a rational influence distinct from the motivational influence of the desire that it’s about.

\* \* \*

Now, if desire-based reasons derive their influence from something other than the desires on which they are based, then perhaps the same influence is available to considerations that aren't based on desires at all. Perhaps considerations that aren't about your inclinations can still provide potentially directive knowledge.

Such considerations would still have the influence of reasons, by virtue of their capacity to engage your inclination toward autonomy. But they wouldn't depend for their influence on the inclinations that differentiate you from any other agent, and they wouldn't be about such inclinations, either. They might therefore be reasons that Williams would call external.

My thesis, in any case, is that reasons for acting shouldn't be classified as external or internal, since they don't conform to the assumptions underwriting the use of these terms. A reason applies only to those whom it can influence, but its application is not therefore limited to agents of a particular temperament. The inclination that makes one susceptible to a reason for acting is just the inclination that makes one an agent.