**Chapter 8: Against Externalism**

At this point it is appropriate to look at some further arguments, positive and negative, in support of internalism. We have already seen one reason why moral realism looks to be incompatible with the thesis: if moral judgments are like other realist beliefs, we should be able to take any attitude we want toward them. In what follows I argue that there are reasons for thinking that our moral realist must in fact be an externalist. If this is so, and if there are good reasons for rejecting externalism (e.g., if we can show that our intuitions heavily favor internalism), then the case against our realist is even stronger.[[1]](#footnote-1)1

We noted at the beginning of Chapter 6 that a great deal rides on the outcome of the debate between the internalist and externalist. If the realist can make a strong case for externalism or show that realism is compatible with internalism, it becomes much more difficult to defend the belief that we should take our affective responses into account when forming our moral judgments.[[2]](#footnote-2)2 In addition, many moral philosophers reject moral realism specifically because they are convinced that it can’t be reconciled with internalist intuitions. Thus, moral realism’s chances of being defensible significantly improve if the realist can show, either that the thesis is compatible with internalism, or that there are good reasons for rejecting our internalist intuitions.

The externalist rejects one or more forms of internalism. The internalist thesis that concerns us is the following:

IT: necessarily, if a person sincerely believes that *she* ought to ϕ (because she believes that ϕ-ing is morally required), then (i) she believes that she has a reason for ϕ-ing and (ii) is therefore motivated to ϕ.

Notice how IT involves the perspective of the agent. The claim is not that the agent necessarily has a reason to ϕ if she believes that she ought to ϕ, or that the agent has a reason to ϕ if morality requires her to ϕ; the claim is rather that the agent necessarily *takes herself* to have a reason to ϕ if she truly believes (and consciously so) that morality requires her to ϕ. When an agent takes herself to have a reason to ϕ, she believes that ϕ-ing can be justified by other beliefs she holds. Notice, too, that the thesis is not ‘. . . if a person sincerely believes that ϕ-ing is morally required, then . . .’. IT concerns only cases in which a person sincerely believes that *she* ought to ϕ.

IT claims that there is a conceptual connection between believing (or judging) that one ought to ϕ and taking oneself to have a reason to ϕ. IT also claims that there is a conceptual connection between believing (or judging) that one ought to ϕ and being motivated to ϕ.

Another way to express IT, or the gist of it, is as follows: necessarily, if a person judges that she ought to ϕ,[[3]](#footnote-3)3 then she is motivated to ϕ. This reformulation aims to make clear that the internalist needn’t be committed to interpreting the judgment “I ought to ϕ” as the expression of a belief state.[[4]](#footnote-4)4 For those internalists who want to maintain a sharp distinction between beliefs and desires, this detail is important. For it is often assumed that an agent won’t be motivated to ϕ unless they have a desire to ϕ. If the judgment “I ought to ϕ” were solely the expression of a belief state (i.e., some state that is entirely non-conative in nature), yet being motivated to ϕ requires the presence of a desire-like state, IT is implausible.[[5]](#footnote-5)5

No one denies that moral judgments have, or can have, practical import. But why think that a *necessary* connection exists between judgment and motivation, between believing (or judging) that one ought to ϕ and being motivated to ϕ? Why think that IT is not only true, but a reflection of some fairly strong intuitions? And why is it so difficult for our realist to hold IT?

I will proceed to answer these questions, first by offering mostly positive support for internalism, then by arguing against externalism by way of denying the possibility of the reliable amoralist. Although much of my argument against the reliable amoralist will be seen by the realist as begging the question, it helps us see why it makes sense for the realist to be an externalist. In fact it is hard to see how any realist committed to the second core tenet could be an internalist. In his book, *The Moral Problem* (1994), Michael Smith argues against this view by trying to show how realism can be made compatible with internalism. Since his argument has drawn widespread attention, I take it up in the final part of this chapter. We see that Smith’s solution to what he calls “the moral problem” isn’t at all persuasive. If anything, his lack of success only adds to our sense that the two theses are irreconcilable.

*Arguments in Support of IT*

It seems right to say that when a person judges that they ought, morally-speaking, to ϕ, they *believe* that they ought to ϕ. It also seems right to say that we can properly attribute moral beliefs to people, and that assertions of the form “I ought to ϕ” express, among other things, a belief of the speaker. We have explored alternative views in earlier chapters of the dissertation but have found them all to be lacking. In what follows, therefore, I will assume that it is appropriate to understand IT as it is first formulated above. But—importantly—I do not assume that our moral judgments are merely expressions of beliefs if beliefs are understood to be entirely non-conative in nature.[[6]](#footnote-6)6

Two questions that arise immediately are: What does a belief about what one ought to do entail? and, What are the conditions for *sincere* belief about what one ought to do?

As the reader saw in Chapter 6, it is my view that the plausibility of (i) largely comes from our sense of the meaning of ‘ought’: to think that someone (including oneself) ought to do something is just to take them to have a reason for doing it. We won’t say that a person consciously and truly believes that she ought to ϕ unless we believe that she takes herself to have a reason to ϕ.[[7]](#footnote-7)7 This seems to be true whatever the sense of the ‘ought’, whether it be understood morally, prudentially, etc.

Some internalists say that IT’s first necessary connection mainly follows from our conception of morality itself.[[8]](#footnote-8)8 When forming moral judgments we consider the kinds of reasons that exist for acting a certain way or, say, for admiring a certain character trait. We are often asking about which actions agents have *most* reason to do, or which character traits agents have *most* reason to develop. If one understands morality along these lines, it seems right to say that believing that one ought to ϕ necessarily involves taking oneself to have a reason to ϕ. We have, then, two very solid reasons for thinking that this first necessary connection holds.

But as obvious as I think the necessary connection is between believing that one ought to ϕ and taking oneself to have a reason to ϕ, the realist will certainly resist the move if we then understand “taking oneself to have a reason to ϕ” in the antirealist manner that I sketched out earlier (in the discussion of feature (c) in Chapter 6). My view is that individuals take something to count as a reason for them (to do or believe something, say) only if it is appropriately related to certain other beliefs and attitudes that they hold; these would be the beliefs and attitudes that constitute that individual’s evaluative, or normative, frameworks.[[9]](#footnote-9)9 So realists and externalists may see even part (i) of IT as controversial; granting the first necessary connection might provide the internalist with too much ammunition for establishing the second.1[[10]](#footnote-10)0

The externalist characteristically claims that a person can sensibly say: “Yes, Y is a moral consideration; I see that it is morally relevant. But even so, I don’t take it to be something I should take into account in my practical deliberations.” Or he claims it is possible for a person to be sincere when they say: “I see that, morally-speaking, I ought to ϕ, but I fail to see that I have a reason to ϕ.”

Yet one needn’t understand “taking oneself to have a reason to ϕ” in antirealist terms in order to think that this view is quite problematic. It is enough to reiterate the reasons already given. First, the view forces us to reject the ordinary meanings of ‘ought’, ‘sincere’, and/or ‘relevant’. Second, it seems to conflict with the common conception of moral inquiry as an inquiry into reasons (of a certain sort) for doing things, or for being a certain kind of person. The externalist will have to explain how it is that we arrive at the moral beliefs we have without relying on the notion of having a reason. This seems to be impossible given how the first part of IT involves the agent’s own perspective. Again, the thesis is not about agents simply taking “do X” to be a requirement of morality; it is instead about the agent who believes that *she* ought to do X because she takes this to be a requirement of morality. It follows from this that the gap the externalist needs for his claim to be true just isn’t there. Here is another way of making this same point:

If, after reflection, you arrive at the conclusion “I ought to ϕ” it must be because of certain considerations. But the considerations which count in favor of ϕ-ing are said to be reasons for ϕ-ing, or the basis for having such reasons. So if you conclude that you ought to ϕ, then it seems this must be because you take yourself to have reasons for ϕ-ing. You would not conclude that *you* ought to ϕ if you thought that what led you to that conclusion could only provide the basis for others having reason to ϕ. If you thought this, what would lead you to conclude that *you* ought to ϕ?

An externalist who is a realist might respond by saying that the judgment “I ought to do X” is a product of direct perception. If moral properties are mind-independent, this is certainly possible. But it means that this externalist must be an intuitionist, and we saw in Chapter 2 why intuitionism is unsatisfactory.

Perhaps some externalists will say, though, that there are other ways for people to end up believing that they ought to do X without having engaged in any reflection. A person might be brainwashed, or indoctrinated, or unwittingly susceptible to the influence of social pressures. Even philosophers don’t have the time to subject all the beliefs they hold to critical scrutiny. At least two things can be said in reply to this suggestion. First, the person who acts unthinkingly, like a machine executing the instructions of a program, is not someone whom we would see as a very good example of a moral agent. One might question whether they are a moral agent at all if they are acting on beliefs which, arguably, they didn’t come to have voluntarily. Second, one can argue that the psychological state which we would identify with this person’s belief that they ought to ϕ is quite different from the psychological state possessed by the person who arrives at this belief reflectively. The states are importantly different, of course, in terms of how they came to be. The argument would then be that this difference has a bearing on whether it is appropriate to attribute the belief “I ought to ϕ” to the agents in question. The success of such an argument will obviously depend on the criteria used to judge whether or not a person does indeed hold some particular belief. But if these criteria are based on the agent’s behavior or on the agent’s inferential and judgmental dispositions (see, for example, (Smith 1994: 30, 37-38), (Smith 1997: 102-103)), it is certainly plausible to think that significant differences will show up between the two agents under different circumstances, due to the fact that the psychological states in question have such different origins. And it is plausible to think that these differences in behavior or dispositions will give us reasons to deny that the brainwashed or indoctrinated or unreflective agent truly believes that they ought to ϕ. Under certain circumstances they won’t think or behave as we would expect them to if they truly believed what they are said to believe—or so one could argue.

For the above reasons, then, I see part (i) of IT as highly intuitive; it simply tells us what a belief about what one ought to do involves, or entails, for the agent holding the belief. I think the same can also be said for part (ii) of IT, although the emphasis will lie, not just with believing something, but *sincerely* or *truly* believing it. Admittedly however—and as the reader will see in what follows—the entailment of part (ii) is less clear-cut.

Whether we might be justified in thinking that a necessary connection exists between sincerely believing that one ought to ϕ and *being motivated* to ϕ obviously depends on how we understand these psychological states. I stick with two commonly held assumptions: first, that one is motivated to ϕ if and only if one desires to ϕ; second, that desiring and believing have opposite “directions of fit”.1[[11]](#footnote-11)1 Again, the claim is that there are intuitions supporting part (ii) of IT which are not obviously antirealist.1[[12]](#footnote-12)2

Let’s start with three preliminary observations.

Our assumption has been, not that an agent believes that she ought to ϕ and then comes to believe that she has reasons for ϕ-ing in virtue of having that belief, but that she sees certain reasons for ϕ-ing and in virtue of those reasons comes to believe that she ought to ϕ. It will be useful then to distinguish between reasons which, for the agent involved, entail the ‘ought’ conclusion and those which don’t. Let’s say that an agent takes herself to have a *decisive* reason, or set of reasons, to ϕ if she concludes that she ought to ϕ and feels at the moment that there is no need for further deliberation. Typically, a reason or set of reasons seen by an agent as decisive will *also* be seen by the agent as overriding the other reasons she may have for doing something that is incompatible with ϕ-ing.

Second, it seems that the most promising way to understand motivation is in terms of dispositions—saying that an agent is motivated to ϕ if and only if she is disposed to ϕ. Dispositional states needn’t always be long-standing ones. Since a motivation to ϕ might be something very specific,1[[13]](#footnote-13)3 we would expect that some motivations can be acquired very suddenly (although we might want to say that this sudden appearance of a disposition is only possible because of certain long-standing dispositions that are less specific in nature). A dispositional account of motivating states seems right as opposed, say, to a behavioral account because, although a person is more likely to ϕ if they are motivated to ϕ, it is not always the case that the agent who is motivated to ϕ actually ϕs. It could be that the agent is more strongly motivated to do X, or that the agent is suffering from weakness of will or depression, or that she faces circumstances that are not conducive to ϕ-ing.

The third observation is that we can hold beliefs with varying degrees of strength. We can believe something but remain skeptical or disposed to give up the belief upon the first receipt of conflicting information. Or, on the opposite end of the scale, we can believe something with firm conviction, so that the weight of conflicting evidence would have to be quite great before we would be willing to relinquish the belief. The degree of our unwillingness to give up a belief in the face of conflicting evidence can depend on a number of things: our view of the epistemic conditions under which we originally came to the belief; the degree to which the belief coheres with our other beliefs; the amount of supporting evidence we have encountered over a period of time; how well the belief has served us since first coming to have it; or how well the belief has withstood critical scrutiny. One further possible source of that unwillingness comes from the effect that firmly held beliefs can have on our dispositions. *Our belief states are not isolated from the other contents of our minds*; this is especially true in the case of our more strongly held beliefs, beliefs which we have relied upon for a long time, beliefs that are in the foreground of our consciousness, and beliefs that we take to be centrally important in our lives and work. Beliefs with one or more of these features are likely to affect the nature and shape of our dispositions. This seems to be particularly true when such beliefs are about what one ought to do, for if we lack the appropriate dispositions, we will likely fail to act upon our beliefs. In any case, the difficulty of relinquishing a strongly held belief may also be due to the need to reshape or relinquish an associated disposition.

These preliminary observations, I shall argue, provide some support for part (ii) of IT without explicitly begging the question against the realist.

If it were the case that a person simply either believes something or they don’t, the use of ‘sincere’ in the statement of IT is redundant. I take it that we don’t see the modifier as superfluous. It won’t be if we think that it tells us something about the *strength* with which the agent holds the belief. The question that concerns us then is: Where does sincerely believing that one ought to ϕ fit on our degrees-of-strength scale?1[[14]](#footnote-14)4

The proponent of IT is claiming that sincerely believing that one ought to ϕ amounts to something *fairly strong*: a person doesn’t sincerely believe that they ought to ϕ unless they are in fact motivated to ϕ. Our three observations help us to see why we find this intuitive. A person *concludes* that she ought to ϕ only if she believes that she has a decisive reason to ϕ. And, unless an agent is engaged in some very technical sort of practical deliberation (the conclusions of which the agent doesn’t intend to act on), decisive reasons will always function like overriding reasons for agents in the sense that decisive reasons are ones agents think they actually ought to act on if they are going to act at all. Thus, the belief, “I ought to ϕ,” will be held fairly strongly. In fact, given the place that moral inquiry has in our lives, we can expect it to be held quite strongly. Very likely it is a product of other firmly held evaluative beliefs, or at least coherent with them. Strongly held beliefs that have a central place in our lives shape our dispositions—in this case, our dispositions to act. And if we are disposed to ϕ, we are motivated to ϕ.

The externalist might try to use the notion that beliefs can be held with varying degrees of strength to his advantage by saying that sincere belief doesn’t entail motivation because one might hold the belief only very weakly. Or, he could insist that it is a mistake to understand sincerely believing something in terms of how strongly the belief is held. A third strategy would be to say that belief doesn’t entail motivation because motivation always involves a desire-like state and beliefs and desires are “distinct existences” (see footnote 5 above). None of these ways of responding, however, are unproblematic.

The problem with the first strategy is that we are not willing to say that a person sincerely believes that *she* ought to ϕ, that she sincerely believes that *she* has a decisive reason to ϕ, if she is not motivated to ϕ. If a person claims to believe, say, that they ought to give to famine relief but then doesn’t actually give when the opportunity is truly there (i.e., when there is no good reason for their not doing so), and if they aren’t suffering from weakness of will or some other psychologically debilitating condition, we conclude that they didn’t *really believe* that they ought to give to famine relief. Either they weren’t sincere when they stated their belief, were sincere but self-deceived, or believed but only in a very weak sense.

The notion that “reasons have motivational implications” (Smith 1994: 6) is even more compelling if we hold a dispositional account of belief states. We might say, for example, that a person sincerely believes that 2 + 2 = 4 if and only if in circumstances C1 they are disposed to make certain inferences or judgments, in circumstances C2 they are disposed to make certain, possibly different, inferences or judgments, and so on. We would expect to see something analogous then for beliefs about what one ought to do: we can say that a person sincerely believes that they ought to ϕ if and only if in circumstances C-1 they are disposed to do such-and-such, in circumstances C-2 they are disposed to do so-and-so, and so forth. (On this account of belief states we can also see what it might mean to hold beliefs with varying degrees of strength. Sometimes one may have only some of the dispositions to infer, judge, etc.) Why limit the dispositions solely to inferential and judgmental ones? Why take dispositions to ϕ to have no connection with the belief that there are decisive reasons for ϕ-ing? Beliefs about how we should live our lives are centrally important to us, are generally strongly held, are certainly much relied upon, and greatly influence what we take to be decisive reasons for doing something. One would expect such beliefs to leave their mark on the nature and shape of an agent’s dispositions to act.

Here the externalist will remind us that he is only denying the existence of a *necessary* connection to motivating states. Admittedly, the proponent of IT doesn’t have a way of *proving* that this connection exists. All she can do is (a) appeal to one’s intuitions about what we take sincere belief to involve; (b) show how IT strongly coheres with the best view of moral properties and moral judgments; and (c) point to the problems we run into if we reject IT. Here I am considering only what the internalist can do in terms of (a) and (c). She claims that the necessary connection holds in virtue of what we take to be involved in holding such a belief sincerely, i.e., in virtue of what we think the meaning of sincere moral belief amounts to. She will also point out that the reason we conclude that the above agent doesn’t sincerely believe they ought to give to famine relief is because we think that this is the best way to make sense of the agent’s behavior.

To this the externalist often responds as follows: ‘There is another way to make sense of the agent: say that she sincerely believes that she ought to give to famine relief but lacks the desire to do so, and this is why she doesn’t give. Sincere belief doesn’t entail motivation because motivation involves having a desire-like state, and beliefs and desires are distinct existences.’

This third strategy may be the externalist’s strongest card. But there are at least three important gaps, or outright difficulties. First, the externalist’s explanation of this agent will need to be supplemented because the explanation as it currently stands implies that whenever agents sincerely believe they ought to ϕ and desire to ϕ, and aren’t suffering from some sort of psychological failure like weakness of will, they will in fact ϕ. Yet it is implausible that the appropriate desire and belief are always sufficient for action in psychologically unconflicted agents. The agent could be caught in what she perceives to be a moral dilemma. Or, I might sincerely believe that I ought to ϕ, desire to ϕ, but then discover that I don’t have the means to ϕ.1[[15]](#footnote-15)5 So, regarding the agent who fails to give to famine relief, there may be other reasons why she fails to give than that she lacks the desire to give. Of course all the externalist needs to show is that it may sometimes be the case that the lack of desire explains the failure to act. Still, it will be very difficult to come up with convincing examples, for we will be inclined to think that an agent doesn’t sincerely believe that she ought to ϕ if she isn’t in fact motivated to ϕ. What the externalist needs is something analogous to Kant’s example of the unsympathetic philanthropist (*Groundwork of the Metaphysics of Morals* 4: 398).1[[16]](#footnote-16)6

Second, the externalist has to explain why we are so often motivated to act in accordance with what we take to be morality’s requirements. Brink says it is simply a contingent fact about us that we are often motivated to act morally (1989: 49). Undoubtedly it is part of our nature to care for other human beings, even perhaps to cooperate with them—seeing that we are social creatures. But saying that the connection is contingent makes it difficult to explain “the phenomenology of moral language use” (Platts 1979: 258), particularly our unwillingness to say that a person sincerely believes they ought to ϕ if they fail to be motivated to ϕ.

Third, the externalist has to offer a defense for the claim that beliefs and desires are always distinct existences. It is entirely implausible to claim that all desires can, in principle, exist in isolation from all beliefs. (One doesn’t desire to go to law school, for example, unless one believes that there are law schools to go to.) Why think, therefore, that our evaluative beliefs or judgments about what one ought to do can exist in isolation from our desires? It is one thing to say that a person can be disposed to ϕ but not believe they ought to ϕ, or that a person can believe that they have a reason to ϕ but not be disposed to ϕ, and quite another to say that a person can sincerely believe that they ought to ϕ and yet not be motivated to ϕ. If we are realists and hold that beliefs and desires are distinct existences, and if we hold that motivation (or caring about morality) involves a desire-like state, then we are committed to the possibility of the “reliable amoralist”—someone who, despite being indifferent to morality’s demands, can reliably arrive at correct moral judgments and competently use moral concepts, even in new cases. Later on in this chapter we see that there are good reasons for denying that such a person exists. If we are antirealists, won’t we already believe that judgments about what one ought to do are evaluations, and that as such, they already involve a desire-like state of the agent? (While there are antirealists who believe that beliefs and desires are distinct existences, I don’t know any of them to be externalists.)

The externalist who claims that sincere belief doesn’t entail motivation because beliefs and desires are distinct existences thus has his work cut out for him. It is not as though he can say that IT is acceptable as long as we understand the modality involved in both necessary connections to be that of logical rather than metaphysical necessity, and then say that his externalism only involves denying the existence of a metaphysically necessary connection between moral beliefs and the motivating states that are typically (from the externalist viewpoint at least) associated with them. In saying this, the thought would be that the phenomenology of moral language use is one thing, reality another; we are simply stipulating that a condition for *sincerely believing* that one ought to ϕ is being disposed to ϕ (and this is why neither side can get very far in the debate without begging the question). The problem with this move is that the externalist contradicts himself if he is claiming that someone can sincerely believe that they ought to ϕ but not be disposed to ϕ.1[[17]](#footnote-17)7 (And if he is not claiming this, what is he claiming that the internalist rejects?) For proponents of IT, including the externalist in question, are committed to the view that a person sincerely believes that she ought to ϕ only if she is disposed to ϕ. So while the externalist who attempts to split the difference is right in thinking that the logical claim (the one interpretation of IT) does not entail the metaphysical claim (the other interpretation of IT), this entailment is not something that the proponents of IT have ever asserted.

Of course, they do intend to be talking about, not just the nature of our concepts of sincere belief and moral judgment, but the psychological states themselves—the nature of our firmly held moral convictions. This leads the externalist to raise the following objection: if splitting the difference is not an option, and yet the proponents of IT are denying that the beliefs and desires in question are distinct existences, then they must be committed to the existence of “besires”—unitary psychological states with opposing directions of fit. The objection is that besires, so-understood, are counterintuitive and highly problematic. There are at least two ways to respond to this objection. First, the metaphor of ‘directions of fit’ needs to be made more precise before we can determine whether besires are truly problematic. Are they as problematic as claiming that beliefs and desires are distinct existences? The claim that one could desire to go to law school without believing that law schools exist, or could exist, is clearly problematic. Second, I am not so sure that the proponent of IT is in fact committed to the existence of besires, precisely because he takes the modalities of IT to be that of logical rather than metaphysical necessity. Since the logical necessity isn’t engaged unless the moral belief in question is sincerely (or firmly, or truly) held, it is not as if the proponent of IT is committed to the idea that moral beliefs, in virtue of their content alone, have a necessary connection to motivating states.

We noted earlier that the externalist might also try another strategy. This second strategy could be to say that we are abusing the meaning of ‘sincere’ by understanding sincere belief in relation to how strongly the belief is held. I think what is really at issue here, however, is whether or not beliefs can in fact be held with varying degrees of strength. For IT could just as well be stated as: ‘if a person *judges* that she ought to do X, then . . .’, or ‘if a person *firmly* (or *truly*) believes that . . .’. ‘Sincere’ itself—this particular word—needn’t be the issue (which is not to say that it is superfluous). Since the externalist will be on shaky ground if he tries to deny that beliefs are held with varying degrees of strength, the second strategy looks to be very weak.

*Psychologically Conflicted, Abnormal, and Practically Irrational Agents*

Those who reject part (ii) of IT often draw attention to depressed or incontinent agents, or sociopaths and amoralists.1[[18]](#footnote-18)8 The claim is that our internalist won’t be able to explain, or won’t have a very good explanation, of the existence of one or more of these kinds of agents. But the proponent of IT can easily account for them, for he never denies that people can fail to desire in accordance with their beliefs. All he claims is that it goes against the meaning of ‘sincerely believes’ to say that a person sincerely believes that they ought to ϕ when they aren’t disposed to ϕ. Also, we have already seen a number of reasons why, under a dispositional account of motivation, the thought that one will always be moved to ϕ if they have a desire to ϕ is quite implausible. Indeed, if it is implausible that the appropriate desire and belief are always sufficient for action in psychologically unconflicted agents, it is even more implausible to think that the appropriate desire and belief are always sufficient for action in psychologically conflicted or impaired agents.

Consider the incontinent person—one who knows what they ought to do, sincerely believes that they ought to do it, but then fails to do it anyway. For example, suppose that our incontinent agent knows that she shouldn’t eat the chocolate cake that is in front of her, sincerely believes that she shouldn’t (because she is diabetic and eating the cake will disrupt her insulin levels), and yet she indulges. According to the proponent of IT, this agent is motivated to avoid eating the cake, and necessarily so. Does the fact that she eats the cake show that she fails to have this disposition? No. One can be motivated to ϕ while also being more strongly motivated, or disposed, to do just the opposite. What is IT’s opponent’s explanation for this agent? How is it better? The explanation, if it is to be proof against part (ii) of IT, must be one that violates (ii) (and yet sounds more plausible). Those who reject (ii) on account of the incontinent agent will thus need to say that the agent doesn’t have any desire at all to refrain from eating the cake. It is certainly possible that the set of reasons for not eating the cake could have nothing to do with her actual desires and dispositions.1[[19]](#footnote-19)9 Because the agent lacks the desire to not eat the cake, this opponent will say, she doesn’t follow through on her belief that she ought not eat the cake. I take it, anyway, that this is the explanation that will be offered for the agent’s incontinence.

It is an explanation that is unsatisfactory in at least three distinct ways. First, it doesn’t say why it is that the agent in question can’t have conflicting desires in the way described by the proponent of IT. (By contrast, I have offered reasons in support of the idea that the person who sincerely believes that they ought to ϕ will be disposed to ϕ.) Second, the explanation, as it currently stands, suggests something implausible: that if the agent has the desire to not indulge, and if she sincerely believes that she ought not to indulge, then she will in fact not eat the cake. Here it is important to keep in mind that when a person suffers from weakness of will it is not that they aren’t motivated to ϕ when they sincerely believe that they ought to ϕ; it is that they don’t ϕ. The latter, and not the lack of motivation, is what distinguishes the incontinent agent from both continent agents and virtuous agents. Third, it is not at all clear how the externalist’s explanation is preferable to that offered by proponents of IT; nor is it clear how it could even be made to be such. For the externalist’s explanation of incontinent agents who sincerely believe that they ought to ϕ but then fail to do so must be unsatisfactory if what I have said earlier is correct, i.e., if we will never grant that the agent in question sincerely believes that they ought to ϕ if they aren’t disposed to ϕ.

What about the depressed agent?2[[20]](#footnote-20)0 Suppose that this agent’s circumstances are the following: she sincerely believes that she ought to ϕ, sees no reason to not ϕ, and yet fails to ϕ. The internalist has no trouble with this agent either. Here again our internalist can say that an agent can be motivated to ϕ and yet fail to do it. Notice that saying this doesn’t force our internalist to reject the widely accepted *belief-desire theory of action*, which says that a complete explanation for any action requires reference to both one (or more) of the agent’s belief-states and one (or more) of the agent’s desire-states. IT is compatible with this theory of action because the latter never claims that certain belief-desire combinations are sufficient for action. The internalist can (and should) explain the depressed agent not doing what she sincerely believes she ought to do, what she is in fact motivated to do and for which she has no conflicting motivations, by pointing to the despair that the agent feels. This despair could be due to brain chemistry and/or due to other beliefs that she holds.

Michael Smith (1994: 135) thinks that we ought to understand the depressed agent as someone who, if she sincerely believes that she ought to ϕ but fails to ϕ, it is because she fails to be motivated to ϕ. The agent fails to ϕ, in other words, because she lacks the desire to ϕ. As we have seen, this suggests the implausible: that being motivated to ϕ and believing that one ought to ϕ are sufficient conditions for ϕ-ing. But even adding the condition, ‘not motivated to do something that conflicts with X’, doesn’t seem enough to ensure that an agent actually does X. One could be motivated to do something that, while not necessarily conflicting with X, takes one’s attention away from doing X. Earlier we also saw other counterexamples to the claim that the appropriate belief and desire set are always sufficient to ensure the action. So even if it is true that the depressed agent might sincerely believe that she ought to ϕ but lack the desire to ϕ, the lack of the desire to ϕ cannot fully explain why she doesn’t ϕ. The empirical evidence, moreover, doesn’t support Smith’s view. The depressed agent’s story is more complicated; she will say, “yes, I know that I should have done X two months ago; I clearly saw the reason for doing it and I desired to do it; I just didn’t do it, and that is (also) what is so depressing.” Geoffrey Sayre-McCord offers a similar criticism of Smith’s position: “Smith thinks the examples of depression and weakness of the will he mentions put the lie to stronger forms of internalism . . . [but] I think the examples Smith cites are not fully compelling, especially once one embraces a dispositional account of motivation, as Smith does. For then, the complete failure actually to be moved to pursue something is compatible with having a desire to pursue it, as long as the disposition that constitutes the desire doesn’t require always being moved to pursue it” (1997: 61, n. 8). Therefore, depressed agents, as a category, also do not tell against the necessary connection to motivation. It is even doubtful whether any particular depressed agent, in virtue of being depressed, is a counterexample to that necessary connection. Again, we run up against the fact that we deny that a person sincerely believes that they ought to ϕ if they aren’t disposed to ϕ.

What about the sociopath? It is said that there are sociopaths who sincerely believe that they ought not to commit their crimes but are not disposed to not commit their crimes; they remain indifferent to moral considerations (Brink 1986: 29). If there are such persons, they qualify as amoralists. But as a proponent of IT, I deny that this sort of amoralist exists. The judge that IT refers to is not arriving at the judgment, “Doing X is the morally right thing to do,” but rather “I ought to do X.” I grant that someone can make the first sort of judgment and yet not be motivated to do X. This is my understanding of the so-called amoralist and the subset of sociopaths that fall into this category: they are capable of recognizing (although not reliably) what others take to be morally correct, but they don’t see that morality applies to them. They can of course utter the words “I ought to do X” but we should take them to be insincere in this utterance; they don’t truly believe that they ought to do X; they don’t truly believe that they have decisive reasons for doing X. Moreover, we should keep in mind that their failure to do X, or their saying that they don’t care about doing X, isn’t proof that they are not disposed to do X; there could be other reasons for their not doing X, or they might merely think that they don’t care about doing X.

I will have more to say about the amoralist in a moment. Before doing so, a final comment is in order regarding the kinds of agents just considered. Often these agents are thought to exhibit one form or another of practical irrationality. Suppose the opponent of IT admits that his use of these agents doesn’t really help to further his position. He might still claim that accepting IT forces us to deny the existence of a whole category of practical irrationality: situations in which an agent sincerely believes that they ought to ϕ but is not motivated to ϕ. The agent is irrational because her desires fail to cohere, or accord, with her sincere belief that she ought to ϕ. In response to this charge the proponent of IT can say, however, that she needn’t deny the existence of this form of practical irrationality; people *can* fail to desire in accordance with their beliefs. The proponent of IT is simply claiming that it goes against the meaning of ‘sincerely believes’ to say that a person sincerely believes that they ought to ϕ when they aren’t disposed to ϕ.

*The Reliable Amoralist*

Some externalists argue, or are said to argue2[[21]](#footnote-21)1, for the possibility of a different kind of amoralist. They claim that it is possible for someone to be a *reliable* judge of morality’s requirements, or to be a competent user of moral concepts, without being motivated to act accordingly. An agent could reliably perceive or recognize moral truths and what is morally relevant without caring a whit about morality. The possibility of this kind of amoralist doesn’t exactly tell against IT—for there is a big jump between judging that ϕ-ing is a requirement of morality and judging that one ought to ϕ because ϕ-ing is morally required—but it does pose a problem for a closely related form of internalism. The possibility of the “reliable amoralist” is a threat to:

IT\*: necessarily, if a person judges that ϕ-ing is morally required, then (i)\* she believes that she has a reason for ϕ-ing and (ii)\* is therefore motivated to ϕ.

The reliable amoralist is actually a category of agent that is much stronger than the externalist needs. All the externalist needs for shedding doubt on IT\* is evidence that a person can make a judgment about what is morally required and not be motivated to act in accordance with that judgment. It is true, however, that we will be much more confident that such a person exists if we think the reliable amoralist is possible.

Should we think that the reliable amoralist directly threatens IT itself? At first glance it appears to. For we generally don’t stand on the “outside” of morality when engaged in practical deliberation. As agents we in fact often view morality in personal terms if we think that moral considerations ought to have any role at all in our practical deliberations. It is often the case, then, that judging that ϕ-ing is morally required is essentially equivalent to concluding that one ought to do ϕ. In this case, IT\* can appear to be saying the same thing as IT.

However, the amoralist is on the “outside” of morality insofar as he doesn’t see it as something that applies to him. He may recognize that in some sense or other (from the standpoint of morality?) he is under certain moral requirements, but this doesn’t matter to him; morality has no practical import for him. Thus the amoralist, I claim, will never be one to conclude “I ought to ϕ” upon recognizing that ϕ-ing is morally required. The amoralist is someone who steps back from morality and questions its authority or relevance; it follows from this act of stepping back that the amoralist cannot sincerely believe that he ought, morally-speaking, to ϕ.2[[22]](#footnote-22)2 If he truly believed this, he could not step back enough to genuinely doubt that morality applies to him, that it is *he* who ought to ϕ.

Therefore, even if it is possible for the reliable amoralist to exist, this amoralist poses no *direct* threat to IT itself. As a proponent of IT, but someone who rejects IT\*, I then maybe needn’t argue against him. But since I am arguing against externalism in order to argue against realism, and since the moral realist is committed to the possibility of the reliable amoralist, it is still worth addressing this type of amoralist. The possibility of this amoralist would be evidence in favor of realism. In this sense, this amoralist is an *indirect* threat to IT (if we assume, as I am, that IT is incompatible with realism).

The moral realist who is an externalist—someone like Brink—will want to say that the reliable amoralist is a real possibility. For if moral facts are truly mind-independent, we should be able to recognize such facts without having to have any particular attitude toward them. If moral facts are mind-independent in the way that we take other facts to be, we should be able to acknowledge their existence but remain indifferent to them; similarly, we should be able to deny that they necessarily have any particular motivational relevance for us. Moral facts or properties look to be queer entities if they are such that they necessarily move us to act, or necessarily give us a reason to act, or if they are such that we must have certain attitudes and beliefs in order to reliably perceive them.

In Chapter 6 I argued that there is a subjective aspect to our moral value judgments that the realist cannot account for due to his commitment to the mind-independence claim. We concluded that moral properties have both objective and subjective aspects to them, aspects which cannot be disentangled. This nondisentanglability leads us to think that properly applying moral terms requires a special cognitive sensitivity. Correctly perceiving moral right and wrong is not just a matter of seeing the external features that might make an action, say, cruel or rude; it is also a matter of knowing when these features do in fact make an action cruel or rude, when they merit the term applied to them. Properly applying moral terms is, in other words, also a matter of knowing when the natural features upon which the terms supervene are morally relevant. In Chapter 6 we saw too that we should take our moral value judgments to be evaluations; we saw that perceiving what is morally right or wrong isn’t simply a matter of what is there but of what we bring to it. If we understand our moral judgments in this evaluative sense, we have a very plausible way to explain how it is that the virtuous agent takes certain features of a situation to be morally relevant but not others. (Moral relevance will be determined by the virtuous agent’s conception of how to live, where this conception is an orectic psychological state.)

The arguments of Chapter 6 led us to conclude that we have to reject the realist’s second core tenet if we are to properly account for the phenomenology of moral value experience. But if we reject the realist’s second core tenet and accept the account of our moral valuations given in Chapter 6, it is difficult to see how the reliable amoralist will be possible. Conversely, if we do try to make room for the possibility of the reliable amoralist, it seems we will no longer be able to adequately account for many of the moral phenomena that we need to account for. Let’s look at this incompatibility in a little more detail.

What follows are three arguments against the reliable amoralist.2[[23]](#footnote-23)3 The first draws upon the nature of reasons and the nature of moral reasons in particular. The second concerns how we judge competency with respect to the use of moral terms. The third regards the uncodifiability of the knowledge needed in order to reliably identify what is morally right or wrong. Each of the arguments reveals a different aspect of the subjective nature of our moral value judgments.

(a) In Chapter 6 we noted that it can be difficult explaining why certain features of situations are morally relevant and not others; or why features which would be morally relevant in some other context are not morally relevant in the present context. Is there some external mark on these features which signals their moral relevance? Or is there some additional, nonsubjective feature that the virtuous agent perceives and which alerts her to the features of a situation that are morally relevant? Neither of these suggestions seems very plausible. The first is untenable because it is possible for the very same feature to be morally relevant in one context but not in another. The second cannot be taken seriously until something more is said about the nature of this additional feature, whether it is the same type of feature across different situations, why we should think it is nonsubjective, etc. In particular, one will have to explain why these additional features are not themselves subject to the fact that “the very same feature can be morally relevant in one context but not in another.”

What is far more plausible is the thought that what enables the virtuous agent to correctly judge moral relevance is the set of concerns and interests that she has come to acquire—what we earlier referred to as the agent’s ‘second nature’ (McDowell 1995b). If we agree with McDowell that the virtuous agent has a distinctive way of seeing things, due to the shape of her second nature, we will be inclined to think that it won’t be possible for someone to reliably recognize morality’s requirements unless their second nature has the right shape, unless the normative framework that they rely on in making moral judgments is of a certain sort.

One reason for thinking that the virtuous agent’s second nature provides her with a distinctive sort of perceptual capacity is the very nature of practical reasons. Since morality is about what one ought to do and the kind of life one ought to live, virtuous agents will be attentive to the kinds of practical reasons there are. The virtuous person acts in the way she does because she takes there to be reasons for acting that way. But importantly, she takes these reasons to be reasons for *her* acting that way. And what she sees as a reason for herself, she sees to be so almost certainly because of the concerns and interests she has (i.e., because of the contents of her subjective motivational set, because of the shape of her second nature). How else could she take what she sees to constitute a reason for her? To accept this line of thought, one doesn’t have to hold that the only kinds of practical reasons are internal reasons (Williams 1980); one just needs to believe that agents *typically* take something to constitute a reason for themselves only if that something is suitably related to one or more elements of their subjective motivational set. For what we are concerned with here is what agents need in order to *reliably* recognize morality’s requirements, i.e., in order to reliably recognize moral reasons.2[[24]](#footnote-24)4

Given this understanding of the nature of (moral) practical reasons, we would expect that the person who doesn’t care about morality, or who seriously doubts whether they have a reason to do what morality demands, won’t have what it takes (especially when confronting new cases) to reliably identify morality’s demands. Identifying these demands is a matter of seeing things in a certain light, a matter of seeing certain *reasons* for acting.

Of course, there are many who reject this understanding of moral reasons. However, denying that the normativity of moral reasons depends in any way on the dispositions and/or desires of the agent seems to commit one to saying that the rationality of virtue is, in principle, demonstrable from an external standpoint. One will also be forced to deny the nondisentanglability of the subjective and objective aspects of moral properties. In other words, rejecting the above understanding of moral reasons in order to allow for the possibility of the reliable amoralist is problematic with respect to the actual phenomenology of our moral value experience.

The amoralist doesn’t see things as the virtuous person does (or even the typical moral agent) because he lacks the normative framework constituted by the virtuous person’s second nature. In asking, “Why should I care about morality?”, the amoralist seeks an external justification for the demands of morality, one which relies only on the normative framework that *he* has. But the nature of practical reasons, and hence of morality’s demands, suggests that such a justification is not to be found.

This, I suspect, is why we agree with Aristotle’s observation that to be a good student of ethics one has to have been properly brought up. The person who is properly brought up has been taught to see things a certain way, and to respond in certain ways to the world. If we are already on the inside of the normative framework created by this proper upbringing, it won’t even occur to us to demand that the requirements of morality be justified by things outside of the framework; nor will it occur to us to ask why it is that we ought to care about morality. Such thoughts won’t arise because (among other reasons) we won’t think that the authority of morality’s requirements comes from outside of us.

By contrast, if the reliable amoralist were possible, the prerequisites for ethics would be far less stringent. The teaching of right and wrong wouldn’t be nearly so difficult, for students of ethics wouldn’t need to have the right set of dispositions or proper upbringing. There might even be adolescent moral geniuses in the same way that there are adolescent mathematical geniuses. But again, all of this goes against what we see and know. We don’t think that such moral geniuses are possible or that moral education is so straightforward. We instead think that Aristotle is right, that being properly brought up is crucial to the project of inquiring into what constitutes a good human life.

(b) The claim is that the amoralist can competently apply moral concepts even though he doesn’t care about morality’s demands, or doubts that they give him any reason to act. But saying this flies in the face of the criteria we use for judging such competence. We don’t think, for example, that a person correctly applies a concept like ‘cruel’ unless they hold a negative attitude toward the object they are calling cruel. As I say in Chapter 6, “correct and sincere use of moral terms requires that the appropriate attitude accompany them.” To deny this is to deny that moral terms have evaluative meanings.

In his “Of the Standard of Taste” Hume remarks on the importance of this noncognitive element for both moral and aesthetic value terms:2[[25]](#footnote-25)5

There are certain terms in every language which import blame, and others praise, and all men who use the same tongue must agree in their application of them. Every voice is united in applauding elegance, propriety, simplicity, spirit in writing, and in blaming fustian, affectation, coldness, and false brilliancy . . . The word *virtue*, with its equivalent in every tongue, implies praise: as that of *vice* does blame; and no man without the most obvious and grossest impropriety could affix reproach to a term which in general acceptation is understood in a good sense or bestow applause where the idiom requires disapprobation . . . . (paragraph 2)

If we agree with the externalist’s claim that the reliable amoralist is possible, it seems that we have to reject Hume’s observation. If this externalist is right, a person can competently apply moral terms without having the noncognitive attitudes which we take to be involved in the very meanings of these terms; we should also be able to *fully understand* moral concepts without referring to the noncognitive features that we always associate with the terms used for those concepts. Yet if moral properties do in fact have a subjective component, this seems very unlikely. If there is an ineliminable subjective aspect to our moral judgments, and thus to our moral concepts, how could such competency be possible?

Hume thought that “there are certain qualities in objects, which are fitted by nature to produce . . . particular feelings” (ibid., paragraph 16). In Chapter 6 I argued that our moral value judgments reflect, or embody, these feelings and that they need to do so. Hume’s notion of a fit between features of objects and human affective responses gets something importantly right about the phenomenology of value experience, for it is not as if the moral, affective responses we have are responses to nothing. And it does seem that if one lacks the particular feeling thought to be appropriate to a situation, one is not perceiving the morally relevant features of that situation as morally relevant. Hume’s notion of fit is just another way of suggesting that the subjective and objective aspects of moral value experience cannot be disentangled. We believe that moral properties supervene on natural features in the world even while we also believe that our moral judgments reflect how things strike us.

If we think that the subjective and objective aspects of moral valuations cannot be disentangled, if we think that we cannot fully understand our experiencing something to be valuable without referring to the kinds of responses we have to that thing, then it is not at all clear how someone who lacks the subjective element will be able to reliably apply moral concepts. Without the appropriate subjective responses, the amoralist will lack a complete understanding of moral concepts, and thus will lack a complete understanding of what is involved in applying them. It is not as if this should come as a surprise. There are other areas besides morality which require us to have certain subjective responses if we are to be competent. For example, it is highly unlikely that someone who is incapable of finding things to be funny could be a good comedian. Similarly, it is highly unlikely that someone who is incapable of experiencing the emotional impact of novels will be good at writing them. The unamused comedian, we think, will have great difficulty judging whether or not certain new material is funny, or might be made to be funny; the unamused comedian will have difficulty even making old material come off as funny because doing so depends so often on the timing of the delivery. Similarly, the unemotional novelist, we think, is unlikely to know how to structure his novels to give them the greatest emotional impact. These beliefs regarding the unemotional novelist and the unamused comedian follow from our belief that perception in such matters isn’t simply a matter of what is there, but of what we bring to it. The beliefs and dispositions, as well as the concerns and interests, of people make a great difference as to what they perceive, and to what they are capable of (reliably) perceiving.

(c) In his paper, “Non-cognitivism and Rule-Following” (1981), McDowell provides us with another argument against the reliable amoralist. As already noted in the material on a second nature, McDowell believes that “we can learn to see the world in terms of some specific set of evaluative classifications, aesthetic or moral, only because our affective and attitudinative propensities are such that we can be brought to care in appropriate ways about the things we learn to see as collected together by the classifications” (1981: 142). Thus,

it seems reasonable to be sceptical about whether the disentangling manoeuvre [i.e., disentangling a cognitive sensitivity from a noncognitive propensity] . . . can always be effected: specifically, about whether, corresponding to any value concept, one can always isolate a genuine feature of the world—by the appropriate standard of genuineness: that is, a feature that is there anyway, independently of anyone’s value experience being as it is—to be that to which competent users of the concept are to be regarded as responding when they use it; that which is left in the world when one peels off the reflection of the appropriate attitude. (1981: 144)

Proponents of the reliable amoralist claim that all one requires for competency in applying moral concepts is the ability to identify the features that are there anyway, the features which competent users of the concept are regarded as responding to when employing the concept. One doesn’t need to care in the appropriate ways about the things being classified.

An important premise in McDowell’s argument is that “A succession of judgments or utterances, to be intelligible as applications of a single concept to different objects, must belong to a practice of going on doing the same thing” (145). He then builds on Wittgenstein’s observations regarding rule-following to conclude that competency in the application of value concepts will require the user of these concepts to see things from the perspective of the “form of life” in which the concepts are employed. What makes this particularly true for value concepts is not only our inability to disentangle the subjective and objective components, but also the uncodifiability of moral principles (due to the complex nature of the question, How should one live?). Since the amoralist is on the “outside” of the normative framework through which moral agents see the world, there is little reason to think that he will have the necessary knowledge to have competency with moral concepts, not if this knowledge is uncodifiable. And we would expect this type of knowledge to be uncodifiable if we think it arises out of an entire conception of how one ought to live, something that is also an orectic psychological state. (See also footnote 57 of Chapter 6.)

Although each of the above arguments against the reliable amoralist might be seen by the realist as begging the question, they give support to our sense that the realist who is an externalist has a great deal of work to do before we have reason to think that his view does a better job of explaining the moral phenomena that need to be explained.

*Why Realism is Incompatible with IT*

In the first part of this chapter I argued that IT is highly intuitive. As such, it is a belief that the moral realist needs to account for if we aren’t to have an additional reason for rejecting his thesis. Or, the moral realist will at the very least need to show that he can account for our internalist intuitions in some other way, that is, when they are expressed differently than what we find in IT. (Of course, whatever this other form of internalism might turn out to be, it too must be highly intuitive.) In the next section of this chapter we look at Michael Smith’s attempt at joining realism with a “mitigated” form of internalism. In this section we want to look at why it is that moral realism is incompatible with IT.

IT is incompatible with any brand of moral realism characterized by core tenet (b), the mind-independence claim. The best way to see this, I think, is by reminding ourselves of why IT\* is incompatible with (b). IT\* says that when a person recognizes and sincerely believes that ϕ-ing is morally required, they are motivated to act accordingly. (b) says that moral facts or truths are independent of our evidence for them—meaning that these moral facts or truths *cannot depend* on the psychological states of the person who comes to recognize them. The problem is that, if a necessary connection exists between believing that ϕ-ing is morally required and being motivated to ϕ, just such a dependence seems to exist. If the realist replies that we shouldn’t confuse the metaphysical status of moral properties with their epistemological status, he will still need to explain why it is that whenever we come to know or believe that ϕ-ing is morally required, we always take ourselves to have a reason to ϕ and are disposed to ϕ. This does not seem to be true of other facts or truths that we take to be mind-independent. Moral properties thus look to be queer relative to other realist entities.

Since IT\* itself is implausible, however, its incompatibility with realism doesn’t directly tell against the thesis. Even so, the realist will have trouble making room for IT because, from the realist’s perspective, IT looks just like IT\*. For the realist, it seems, the two distinct forms of internalism will have to be equivalent. There are at least two reasons for thinking this. First, for those of us who are not amoralists, IT and IT\* are equivalent in practice. When we are on the “inside” of morality and judge that ϕ-ing is morally required, we are concluding that ϕ-ing is morally required for us—that we ought to ϕ. For the realist who is an internalist, *everyone* will always be on the inside of morality—no special perspective will be needed in order to recognize moral properties or in order to be motivated to act morally. Second, it would seem that for the realist the belief-state “I ought, morally-speaking, to ϕ” must essentially be the same as the belief-state “ϕ-ing is morally required” because the normative authority of realist moral facts is categorical in nature; and because if the realist allows that there is a difference between these two belief-states, it may be difficult to defend it without violating other realist commitments. The realist interprets “I ought to ϕ” as an assertion that says something about how things are in the world, presumably that ϕ-ing is morally required. The truth that is perceived is not that ϕ-ing is morally required only for the agent perceiving it. Realists don’t usually reject the universality constraint2[[26]](#footnote-26)6 or the categorical nature of moral requirements; moral facts are said to be objective and non-relative.

But even if the realist interprets the contents of the belief state expressed by the words “I ought to ϕ” in a way that is specific to the perceiver, and so is maybe in a better position to say why there is a connection between this belief state and a motivating state of the agent, he still has to explain this connection to motivation without contradicting the mind-independence claim. How could there be a *necessary* connection between what is purely a belief state, whose content represents the existence of a moral fact in the world, and motivating states of the agent? Since this belief state will sometimes be caused by the perception of the moral fact (this will be so, I presume, in some cases at least; in other cases it may be that the belief state largely comes about through reasoning and reflection2[[27]](#footnote-27)7), we have the appearance of a necessary connection between mind-independent moral facts existing in the external world and motivating states of agents. If the realist is a rationalist,2[[28]](#footnote-28)8 one who claims that beliefs alone can move us to act, then he is also committed to the implausible view that rational argumentation alone is, in principle, always capable of motivating others to act in accordance with the demands of morality. And the realist who is a rationalist still has to explain why we ought to think that the connection is a necessary one.

The moral realist has problems with IT, then, because IT is still about a necessary connection between certain belief states and certain motivating states. No such necessary connection holds if motivating states involve desire-like states and if beliefs and desires are distinct existences. But claiming that there is a necessary connection between the two kinds of mental states, and thus claiming that beliefs and desires are not distinct existences, makes it difficult, if not impossible, to hold the mind-independence claim, assuming we understand judgments like “I ought to ϕ” in typical realist fashion.

*Why we ought to think that the moral realist won’t be able to incorporate our internalist intuitions in some other way*

Michael Smith claims to have resolved what he calls “the moral problem”, the problem of combining “the objectivity of moral judgments” (our belief that the correctness of moral judgments is wholly determined by external facts) with “the practicality of moral judgments” (our belief that correct moral judgments reflect the reasons we have for doing things) (Smith 1994: Chapter 1). Smith argues that the moral realist can be an internalist. If Smith is right, my arguments to the contrary are somehow flawed. So it is worth taking a brief look at his project.

As Smith and others have suggested (e.g., (McNaughton 1988: section 2.3)), the central problem which the realist and his opponents are trying to grapple with is that of making room, in their understanding of the nature of morality, for three key beliefs. We believe that 2[[29]](#footnote-29)9

(1) Moral judgments of the form ‘It is right that I φ’ express a subject’s beliefs about an objective matter of fact, a fact about what it is right for her to do.

(2) If someone judges that it is right that she φs then, *ceteris paribus*, she is motivated to φ.

(3) An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume’s terms, distinct existences.

These three beliefs appear to be inconsistent, and this is where the problem lies. Smith writes:

The apparent inconsistency can be brought out as follows: from (1), the state expressed by a moral judgment is a belief, which, from (2), is necessarily connected in some way with motivation; that is, from (3), with having a desire. So (1), (2), and (3) together entail that there is some sort of necessary connection between distinct existences: moral belief and desire. But (3) tells us that there is no such connection. Believing some state of the world obtains is one thing, what I desire to do given that belief is quite another. (Smith 1994: 12)

One way realists have tried to resolve this problem is by giving up the belief-desire theory of action explanation, that is, (3). But the realist who rejects this theory then needs to offer another, equally plausible theory of action—something which isn’t easy to do. Because the function of a belief is to represent or convey how things are in the world, it doesn’t seem that beliefs alone can motivate an agent to action. But desires alone also fail to give a complete explanation of an action since, though they move us to act, they don’t seem capable of telling us why an agent specifically does what she does.3[[30]](#footnote-30)0 Nor does it seem (others have said) that we can plausibly explain an action in terms of some other single state that is a combination of both a desire and a belief (a “*besire*”), for beliefs and desires, as we presently understand them, have exactly opposite “directions of fit” to the world. Because we want our beliefs to accurately represent the way things are in the world, we change them, not the world, in order to make the two fit. With our desires, however, things are just the opposite: if we find that the world does not fit the aim of our desires, we seek to change the world, not our desires. How could a single state contain such opposite characteristics? If the world doesn’t fit this single state, the besire, do we change the besire to fit the world or do we attempt to change the world to fit the besire?

Antirealists often try to resolve the moral problem by giving up (1), but we saw in Chapters 4 and 5 the kind of difficulty this leads to.

Smith’s solution is to say we can be realists and hold all of (1), (2), and (3). The way the realist can do this, he maintains, is by distinguishing between *motivating* reasons for action and *normative* reasons for action:

Normative reasons are considerations, or facts, that rationally justify certain sorts of choices or actions on an agent’s behalf. They are propositions of the form, “Acting in such-and-such a way in so-and-so circumstances is desirable.” Motivating reasons, on the other hand, are psychological states with the potential to explain an agent’s action teleologically, and perhaps also causally. They are pairs comprising desires and means-end beliefs. (Smith 1997: 87)

An agent can have a normative reason to do X regardless of the contents of her mind (Smith 1994: 96).

Smith says that we can resolve the moral problem by understanding the Humean theory of motivation, what is the essence of (3), as a theory about motivating reasons, not normative ones; and by understanding our judgments about moral right and wrong as expressions of our beliefs about the normative reasons we have. With this distinction in hand, Smith argues that we can see, by way of conceptual analysis, that a belief about what one has a normative reason to do *just is* a belief about what one would want oneself to do if one were fully rational. Add to this the crucial premise that a person is not fully practically rational unless their desires cohere with their evaluative beliefs (C2 on p. 148 of Smith 1994) and we can argue as follows to show why it is that the realist can hold (2): if a person judges that it is right that she φs in C, she believes that she has a normative reason to φ in C; she believes, in other words, that she would want to φ in C if she were fully rational. So, if she *is* fully rational, she will indeed want to φ in C, and necessarily so—by the very definition of ‘fully practically rational’. This is the extent of the necessary connection of (2), the extent to which we should commit to internalism, Smith says. (In other words, Smith rejects both IT and IT\* and would say that, while my argument that the realist cannot espouse IT may be sound, it is a mistake to think that this tells against moral realism.) Finally, we can hold (1) because we understand moral judgments to be beliefs about the normative reasons that we have. We solve the moral problem, then, by being Humeans in our theory of motivation, our understanding of motivating reasons, but anti-Humeans in our understanding of normative reasons (for the Humean, it is said, denies that desires can be the objects of rational criticism).

The above is the barest outline of Smith’s overall argument. It is worth emphasizing, as Sayre-McCord does in his assessment of that argument, that Smith tries to reconcile (1) - (3) by defending “an account of our moral concepts”. This account is a conceptual analysis of those concepts (Smith claims), one which “promises to show both that in deploying the concepts we are expressing beliefs about objective matters of fact and that their deployment nonetheless bears an appropriate necessary connection to one’s motives, all without contravening the assumptions of Humean psychology” (Sayre-McCord 1997: 57). Smith’s argument is very lengthy and complex; here is not the place to look at it in all its details. I do want to look at two things in more detail, however, before stating some criticisms. One is the nature of normative reasons, the other is the nature of moral facts.

Smith argues that it is part of the very concept of a normative reason and of being rational “that fully rational people converge in their desires” (1994: 166 ff.). It is not part of our conception of a reason for action that such reasons are relative to individuals (167-173). When you and I disagree over reasons for acting, we are having a genuine disagreement; we are talking about reasons *simpliciter*, not about reasons-for-me and reasons-for-you (167). So we must recognize that we have a non-relative conception of normative reasons, one wherein “everyone can reason themselves to the same desires if they engage in a process of systematic justification of their desires. Which desires *I* would end up with, after engaging in such a process, thus in no way depends on what *my* actual desires are to begin with. Reason itself determines the content of our fully rational desires, not the arbitrary fact that we have the actual desires that we have” (Smith 1994: 173). None of this is to say that such non-relative normative reasons exist (173). But Smith argues in the last chapter of his book that we have good reason for thinking that there are reasons that we all share. In fact, to deny the existence of such reasons is to deny that there are moral truths. So Smith’s form of moral realism depends on the existence of these non-relative normative reasons for acting.

The convergence in desires that Smith speaks of is, he says,

very circumscribed. There is no suggestion that fully rational people will all have the same tastes in food, and clothes, and basketball teams. On the contrary, they will presumably be at least as culturally and individually diverse as many human beings throughout history have been. The claim is rather that they will all converge in their desires about what is to be done in highly specific circumstances. Characterize a choice situation in its entirety . . . and, I say, fully rational creatures will all converge on a desire that the very same course of action be pursued. (Smith 1997: 89)

Smith also tells us what his view means for the nature of moral facts or truths, and for the nature of moral properties. If I understand Smith correctly, moral facts or truths just are certain kinds of normative reasons. Normative reasons themselves “are . . . best thought of as truths: that is, propositions of the general form ‘A’s φ-ing is desirable or required’” (Smith 1994: 95). If a moral judgment that we make happens to be correct, and if, as Smith suggests, it is right to say that moral facts are expressed by correct moral judgments, then moral facts are facts about what we have normative reason(s) to do (1994: 59, 185; 1997: 117); and “what we have normative reason to do is what we would desire that we do if we were fully rational (1994: 150). “Facts about what we have normative reason to do . . . are facts about the desires we would all converge on if we were to come up with a maximally informed and coherent and unified set of desires” (Smith 1997: 97). Asking whether there are moral facts is equivalent to asking whether there are any normative reasons for acting (1997: 116); “A normative reason to maximize happiness and minimize suffering . . . [is] an excellent candidate for a moral fact”; so too are the “normative reasons corresponding in their content to [W. D.] Ross’s seven prima facie duties” (Smith 1997: 117). Moral facts or truths give us normative reasons for acting, then, in the sense that such facts or truths just are normative reasons.

We are also told that because normative reasons are non-relative, they apply to agents categorically, not hypothetically (Smith 1994: 175). Moral truths are necessary, a priori truths (ibid., 187, 191-192). Moral properties, if there are any, just are natural properties (1994: 57). Yet since Smith espouses a non-reductive (or non-definitional) form of naturalism, we cannot translate moral, evaluative terms into nonmoral, nonevaluative terms (1994: 185-86). He gives us the following example in support of his claim that his moral realism is a form of naturalism: “the rightness of acts in certain circumstances C . . . is the feature that we would want acts to have in C if we were fully rational, where these wants have the appropriate content” (ibid.). But if we should discover, substantively, that Fness is the feature that we would want acts to have in C if we were fully rational (what in Smith’s view is a necessary, a priori truth), we can say that “Rightness in C is Fness” (ibid.).3[[31]](#footnote-31)1 This is the sense, Smith says, in which his moral realism “squares with” a “broader naturalism” (ibid.).

Does Smith’s solution undermine my contention that the moral realist must be an externalist, particularly one who rejects IT? I think we can confidently say ‘no’, and for the following reasons.3[[32]](#footnote-32)2

It is important to recognize, first of all, how it is that Smith has gone about reconciling (1), (2), and (3). Despite what he claims, he gives up the very essence of (3); he seriously waters down (2); and it is not at all clear that he holds (1) in a way that is compatible with the realist’s second core tenet. Let’s look at this set of problems first.

The internalism that Smith espouses is extremely weak. Let’s follow Sayre-McCord (1997) and call it “mitigated internalism”:

**MI**: If an agent judges that it is right for her to φ in circumstances C, then either she is motivated to φ in C or she is practically irrational (Smith 1994: 61).

It would not be unreasonable or inaccurate to describe MI as a form of externalism, especially if Smith attributes the existence of the necessary connection entirely to our conception of the fully practically rational (hereafter ‘FPR’) agent, or if, as I think is the case, he attributes the existence of the necessary connection entirely to our conceptions of the FPR agent and normative reasons—as opposed to saying that the necessary connection is due in part to the nature of the belief states involved. Smith argues that it is a mistake to subscribe to any stronger form of internalism; depressed agents in particular, he thinks, tell against IT. But we have seen that Smith’s view of the depressed agent is questionable. We have seen, in other words, that Smith does not have a good argument against stronger forms of internalism.3[[33]](#footnote-33)3

As Smith understands (3), or presents his understanding of it, there are three separate claims involved.3[[34]](#footnote-34)4 The first is that beliefs and desires are distinct existences. The second is the belief-desire theory of action. The third is the Humean theory of motivation—what Smith understands as the claim that motivating reasons are *constituted of* appropriately related beliefs and desires. In his discussion of this theory of motivation (1994: 92-93; 1987: 59), Smith *suggests*, and rightly so, that the Humean is committed to the *desire-out, desire-in principle*; this is roughly the idea that the existence of a desire can never be due to some non-desire state that is rationally justifiable; rational justification of a desire (if it should admit of any such justification at all) can never terminate in something for which rational justification can be given; what must be cited, in the end, is another desire (see (Wallace 1990: 370)).3[[35]](#footnote-35)5 It follows from this principle that an agent’s ends cannot be rationally criticized; the scope of practical reason “will be restricted to accounting for the extension of motivational influence from given, antecedent ends” (ibid.).

The *rationalist* is someone who takes an opposite view. The rationalist will say, contra Hume, that reason alone is capable of moving us to act, that “the psychological processes which originally give rise to motivation can be processes which are governed . . . by the principles or norms of reason” (Wallace 1990: 355). It is important to note, however, that this rationalist thesis is still compatible with the belief-desire theory of action (something that Smith recognizes; see (1994: 93)). For the latter says nothing about *the source* of the desire states in question, about how they came into being. The rationalist who holds this theory of action only needs to allow that some of our belief states can account for the existence of certain desire states. He can then understand the belief-desire theory of action as saying only that both a belief and a desire must be *present* in the agent in order to explain the agent’s doing X (where it is assumed that the belief and desire that are present are both appropriately related to the action of doing X).

Surprisingly, Smith ends up rejecting the Humean theory of motivation (or what, as Wallace so cogently argues, is the real essence of it):

The truth of C2 [“If an agent believes that she has a normative reason to φ, then she rationally should desire to φ” (p. 148)] has obvious repercussions for the nature of deliberation. Suppose an agent who does not yet desire to φ deliberates and, as a result, comes to believe that she has a normative reason to φ. And suppose further that her coming to have this belief then causes her to desire to φ. Given C2 it follows that we should redescribe this causal transition between belief and desire in normative terms. For her having that belief causes her to have a desire that it is rational for her to have, given her belief . . . .

Moreover, the fact that our beliefs and desires may bear such normative relations to each other is not inconsistent with the Humean theory of motivating reasons defended earlier. Indeed, this whole discussion has been premised on the Humean theory. All actions are indeed produced by desires, just as the Humean says; no actions are produced by beliefs alone or by besires. But, if what we have said here is right, some of these desires are themselves produced by the agent’s beliefs about the reasons she has, beliefs she acquires through rational deliberation. (Smith 1994: 179)

Smith turns out to be a rationalist after all, for he allows that certain belief states can bring about motivating states. Of course, if one gives up the Humean theory of motivation, one’s commitment to the belief-desire theory of action isn’t all that significant any more. What does it matter that belief states cannot constitute motivating reasons if they can cause such reasons to exist (see footnote 37)?

While in one sense it is quite surprising to find that Smith is a rationalist,3[[36]](#footnote-36)6 in another sense it is exactly what we would expect given that he holds (1), is also trying to hold (2), takes beliefs and desires to have opposite directions of fit, and insists that motivating states are constituted in part by desires.3[[37]](#footnote-37)7 If moral judgments are truly expressions of belief states, (2) asserts the existence of a necessary connection between certain belief states and certain desire states. The only way to establish this connection is by being a rationalist; the Humean cannot hold both (1) and (2), not even if (2) is watered down into MI.

It should be noted that nothing of what has been said thus far undermines Smith’s solution to the moral problem. The only problem we have encountered is that Smith claims to be a Humean with respect to motivation when he is not.3[[38]](#footnote-38)8

What about Smith’s insistence that he holds (1) as the realist does? In Chapter 2 I noted that Smith characterizes the moral realist in different terms than Brink does (see also (Smith 2000a: 15). It is possible that Smith thinks the moral realist is *not* committed to the mind-independence claim. Yet, as I argued in Chapter 2, it is not at all clear that a view is realist if it doesn’t include the second core tenet. Moral realism is a metaphysical thesis; as such, the realist is understood to be making an additional metaphysical commitment relative to nonrealists. By itself, the first core tenet doesn’t involve such an additional commitment. So if Smith rejects the second core tenet but espouses the first, we should see him as a constructivist, not a realist. If he *is* a constructivist, his solution to the moral problem doesn’t undermine my claim that the realist must be an externalist.

While Smith doesn’t explicitly say that his realism includes the mind-independence claim, there is some evidence that his form of realism is committed to it.3[[39]](#footnote-39)9 But in that case he needs to argue for the claim that *the realist* can consistently say that if we take X to be a moral requirement, we necessarily take ourselves to have a normative reason to do X (1994: 62, 200). He cannot simply assume that this internalist thesis is compatible with the realist’s mind-independence claim.4[[40]](#footnote-40)0

Moreover, if Smith’s realist is committed to the second core tenet, Smith has to explain why it is that moral facts are unlike other mind-independent facts; for Smith’s moral facts are such that we cannot take any attitude we want toward them, not if we are fully rational. Smith does offer a partial explanation: he tells us that moral truths are a kind of normative reason, and proper conceptual analysis of ‘normative reason’ and ‘fully practically rational’ shows us why FPR agents will be disposed to do what morality requires. Unfortunately, this conceptual analysis doesn’t answer the worry about mind-dependence. Let’s look at why this is so.

In that analysis we are told that moral facts are “facts about the desires we would all converge on if we were to come up with a maximally informed and coherent and unified set of desires” (Smith 1997: 97). Moral facts are facts about what an Ideal Advisor would want us to do in the situations we face, where this Ideal Advisor has a subjective motivational set that is *not* radically different from that of the agent getting the advice. This last condition is needed because Smith holds that all normative reasons are ‘internal reasons’. For Smith, normative reasons aren’t reasons for an agent unless they can somehow be linked with the elements of that agent’s subjective motivational set. (See (Kirchin 2005) and (Smith 1995).)4[[41]](#footnote-41)1 We saw earlier in our discussion of well-being why this understanding of moral facts might be attractive. Our responses to things and the desires we have are importantly relevant to our well-being. Indeed, the value or disvalue that things have for us (even when considered from an ideal perspective of being fully informed, etc.) is oftentimes largely a function of how we respond to them—how desirable or undesirable we find those things to be. But if moral facts are facts about the desires FPR agents would converge on, Smith has to say why it is that moral facts are mind-independent, especially since the desires of the Ideal Advisor cannot be radically different from our own.

Recall that for the realist *what makes* a moral judgment true is something other than the beliefs or attitudes we have—whether as individuals or as groups of individuals—regarding the object of the judgment. While Smith may be right in saying that the desires we would all converge on *as FPR agents* will not depend on the actual desires of any *particular* agent who engages in the process of systematic justification of their desires, Smith himself recognizes that the desires *we* converge on *in this process* will be a function of the desires of all those who engage in the process (1994: 175-76). In fact, it is likely that there won’t be any convergence at all unless all those engaged in the project of systematic justification can agree on the starting points for this project. But Smith sees this only as an epistemological difficulty, and thus not as something that reflects on the actual nature of moral facts themselves. If there are any such facts, there will be desires that FPR agents converge on. Still, why should we think that these facts are mind-independent? Our ability to know about them depends on the desires that we have, perhaps not as individuals, but as groups of individuals. Smith claims, however, that we are in the same position, epistemically-speaking, with respect to other sorts of facts, such as facts about what supports what (176); our scientific theories are dependent on starting-points in exactly the same way (201). But then, maybe we should be antirealists with respect to our scientific claims and logical truths.

For Smith, there must be something about the desires of FPR agents which allows us to say that these desires are entirely independent of the desires of real people. Otherwise moral facts won’t be mind-independent. Of course, if Smith’s moral facts truly have such mind-independence, they would seem to lose the appeal that we spoke of earlier, for it seems that they would no longer be relevant to our conception of human well-being. In any case, what will get us the needed independence?

Smith, I think, will say the following: we need to distinguish between rational and non-rational desires (or, we might say, between motivated desires and unmotivated desires; see (Wallace 1990)). Rational desires are those for which justification can be provided. Among the rational desires themselves, we should distinguish between those which have moral content and those which do not. The rational moral desires will be ones which accord with the requirements of morality (or more generally, the norms of practical rationality). These rational moral desires are the ones that FPR agents converge on. Since the requirements of morality (or norms of practical rationality) determine the desires and justify them (Smith defines the FPR agent to have no false beliefs and all relevant true beliefs), and since the former are non-relative, the desires will be non-relative too. If the requirements of morality indeed determine the desires, we will have the objectivity and mind-independence that realism requires.

There are a couple of problems with this response. First of all, it *assumes* that moral facts are mind-independent when it needs to give us reasons for thinking this, especially since moral facts are analyzed in terms of desires. According to Smith, the normative reasons that constitute moral facts necessarily have motivational force for fully rational agents. Second, the desires that FPR agents converge on must be desires that real agents can come to have by deliberating from the elements that are already in their subjective motivational sets, for Smith holds that the only kinds of reasons there are are internal ones. The above response doesn’t explain how the desires that FPR agents converge on can both be a function of the desires that people actually have, and yet not be a function of them.

Smith defends the claim that moral facts are mind-independent, or non-relative, by arguing that this is what analysis of the concept of a normative reason shows us (1994: 166-174). When we argue over a claim like “It is desirable that p in circumstances C”, we assume that the claim “has a straightforward truth condition” (166). We won’t see ourselves as having a genuine disagreement if we didn’t make this assumption. And if, Smith claims, we don’t take the reasons we offer to be non-relative, we potentially end up talking past each other. One person could be arguing about whether from the perspective of person A it is desirable that p in C, whereas the other person could be arguing about whether from the perspective of B it is desirable that p in C. Although Smith makes several further observations in support of his claim that normative reasons are nonrelative, I don’t find the argument at all convincing. I see at least four problems.

First, Smith assumes that assertions about what it is desirable to do are either completely non-relative or relative to individual agents. But there is at least one other important possibility. Perhaps normative reasons for acting are relative to groups of people. Smith offers no reasons for thinking that moral requirements are not culturally relative or relative to different forms of human life. Two people working with or in the same normative framework can have a genuine disagreement over claims about the desirability of things. And, if we recognize that reasons for acting are, or might be, relative in some way, we adjust the ways we argue over such reasons accordingly, exactly so that we don’t end up talking past each other when trying to settle a dispute.

Second, Smith’s argument completely ignores all the beliefs and dispositions which contradict his view of normative reasons. Given his analysis, how would we explain the existence of the tension discussed in the first pages of Chapter 1 (see also section 1.1 of (McNaughton 1988))? Smith ignores, in particular, the implications of his belief that there are no external reasons. If the only reasons are internal ones, then normative reasons must somehow be a function of the desires that agents actually have. But if the desires that we start with when engaging in the project of systematic justification of our desires make a difference to the outcome of that project, why won’t this also be true for FPR agents? In saying that FPR agents won’t all have the same tastes in food, clothes, or basketball teams, in saying that there will be cultural diversity, etc. among FPR agents, Smith seems to assume that FPR agents can isolate moral values from everything else that they value in life. Such isolation seems not only unfeasible but undesirable. It also seems unlikely, since FPR agents will have maximally coherent and unified sets of desires. If FPR agents show wide-ranging differences in these other areas of value, and if these other areas of value impact one’s moral values—as one would expect given that morality is about how one ought to live, and given that FPR agents have maximally coherent and unified sets of desires—, then there is good reason to think that the morally relevant desires of FPR agents won’t converge. I think we even see this understanding reflected in our conception of a normative reason—which brings us to a third point.

Third, it is simply not true that we undisputedly conceive of normative reasons as nonrelative. If we do conceive of normative reasons in this way, then almost everything I said in Chapter 6 about normative frameworks is incorrect. Smith himself writes:

To say that someone has a normative reason to φ is to say that there is some normative requirement that she φs, and is thus to say that her φ-ing is justified *from the perspective of the normative system* that generates that requirement. . . . normative reasons are . . . best thought of as truths . . . These truths may well be many and varied, as many and varied as there are normative systems for generating requirements. Thus, for example, there may be normative reasons of rationality, prudence, morality, and perhaps even normative reasons of other kinds as well. (1994: 95-96; my emphasis)

Smith goes on to try and show that the normative requirements of morality just are normative requirements of rationality. But even if we accept his argument for this view, it is unlikely that many people will say that moral requirements are coextensive with prudential requirements. If we can manage talking about what we ought to do, prudentially-speaking, versus what we ought to do, morally-speaking, why can’t there be different normative frameworks for morality itself? Why isn’t it possible for there to be different forms of life for human beings?4[[42]](#footnote-42)2

Fourth, in arguing for the nonrelativity of normative reasons, Smith suggests that we are committed to saying that rational justification itself is relative if we believe that normative reasons are relative (1994: 171). But this is to say that the norms of practical rationality, supposing that there is just one set of norms, can completely determine an ethical life for us. Why think that these norms could generate all the substantive content of morality, especially given the specificity of many first-order moral questions? Here I think Smith faces the same kind of problem that Williams identifies for forms of Aristotelian naturalism (Williams 1985: Chapter 3). Williams writes: “it is hard to believe that an account of human nature . . . will adequately determine one kind of ethical life as against another” (52). Substitute ‘rationality’ for ‘human nature’ and we have the worry I want to express for Smith’s project. If there are reasons which we all share just in virtue of being human (as there needs to be in order for there to be the moral facts that Smith needs), why think this set of reasons will be extensive enough to determine one kind of ethical life as against another?

Let’s return for a moment to the fact that Smith never addresses the conflict that exists between his claim that normative reasons are nonrelative and his claim that normative reasons are internal ones.4[[43]](#footnote-43)3 In his analysis of normative reasons, Smith ignores that he is committed to saying that the desires which FPR agents converge on must be a function of the desires that are already present in individuals’ subjective motivational sets if those individuals are ever to be able to feel the normative force of moral requirements.4[[44]](#footnote-44)4 Given this conflict, it is worth asking whether Smith shouldn’t give up his commitment to internal reasons.

Is it possible for Smith to hold (1), (2), and (3) and believe that there are external reasons? Smith is committed to the belief-desire theory of action and to the idea that moral judgments are expressions of belief states. He also holds that there is a necessary connection between judgment and motivation in the FPR agent, and that motivating states are constituted in part by desires. Desires belong to an agent’s subjective motivational set. The FPR agent has no false beliefs and all relevant true beliefs; from this fact it follows that if normative reasons could be external reasons, the FPR agent could have moral beliefs without being motivated to act in accordance with them. This, in turn, would mean that Smith has to give up his version of (2).

If the normative reasons that constitute moral truths are external reasons, then something can be a moral requirement for agents without it being deliberatively accessible to them from their current subjective motivational set.4[[45]](#footnote-45)5 This means, however, that rational argumentation alone won’t suffice to get agents for whom a moral truth is an external reason to be motivated to act in accordance with this moral truth. It means, in other words, that Smith’s version of internalism is false. Smith thinks that this contradicts “common-sense assumptions about the nature of morality” (1999: 43). He believes that “the task of getting someone morally motivated [is] no different to the task of getting her to believe what is true” (ibid.). If we deny this, we are forced to think that getting others to change their desires or dispositions will have to involve “a process akin to *conversion*” (ibid.); we are forced to think that “the task of getting people morally motivated [is] no different to the task of getting them to buy this or that product as the result of a cleverly devised and manipulative advertising campaign” (1999: 44). For Smith, therefore, the rationality of virtue is demonstrable from an external standpoint; there is no being on the ‘outside’ of morality. The possibility of moral requirements being external reasons would thus also conflict with Smith’s commitment to rationalism.

In the last several pages I have tried to show why we have reason to doubt that Smith can hold (1) as the realist does. Smith argues that the mind-independence of his moral facts follows from an analysis of normative reasons, but we have seen several reasons for thinking that his analysis is inadequate. We have seen, too, that Smith gives up the Humean theory of motivation, what is the heart and soul of (3). What about (2), or Smith’s form of internalism? Is Smith’s argument for MI sound? Geoffrey Sayre-McCord (1997) shows decisively that it is not.

Sayre-McCord’s argument begins with an excellent summary of Smith’s project. The elements of this summary are important enough to the argument that they are worth setting out. He notes that Smith’s account of moral concepts

comes in three stages. The first stage is to defend a version of moral rationalism according to which judgments about *what we are morally required to do* are judgments about what we have reason to do [cf. (Smith 1994: 62)] . . . The second stage is to defend the view that judgments about *what we have reason to do* are judgments about what we would want ourselves, as we actually are, to do if we were fully rational [cf. (Smith 1994: 150-155)]. And the third stage is to spell out *what it would be for us to be fully rational* in terms of our having no false beliefs, our having all relevant true beliefs, and our having deliberated fully and correctly (where this involves, among other things, our engaging our imagination appropriately and our having acquired a maximally coherent “desiderative profile”) [cf. (Smith 1994: 155-161)].

The end result is an Ideal Advisor theory of reasons and, derivatively, moral rightness. Thus, as Smith would have it, a person has reason to φ in C if and only if were she fully rational she would want herself, as she actually is, to do it . . . It is morally right for a person to φ in C, Smith then holds, if and only if she has a reason (of the appropriate sort and strength) to do it. (Sayre-McCord 1997: 58)

Sayre-McCord then notes that MI commits Smith to the view that “one is always and necessarily irrational if one fails to be motivated to do what one believes to be morally right” (1997: 61). To defend MI, Smith relies on the following claim (which he labels ‘C2’):

C2: If a person believes she has a reason to φ, then she rationally should desire to φ.4[[46]](#footnote-46)6

Sayre-McCord observes that we have to understand “rationally should” in a certain way if we are to think that this statement is true. It is not the case that a person *has a reason to* desire to φ just because she believes she has a reason to φ; her belief could be false. Even if her belief is true, it is sometimes the case that we cannot desire to φ if we are to succeed at φ-ing (Sayre-McCord 1997: 63). We can make better sense of C2 if we translate it as “If a person believes she has a reason to φ, then *if she is rational* she will desire to φ” (ibid., 64; my emphasis). As we have seen, this is exactly how Smith wants us to understand C2, for according to Smith, when a person believes that she has a reason to φ, she believes that if she were fully rational she would want herself to φ. Smith maintains that a person is not fully practically rational unless she desires to do what she believes she has a reason to do.

But this understanding of C2 leads to a problem for Smith because “we now have two standards of rationality in play—one satisfied when one does (or desires, or thinks, or whatever) as one has reason to, the other satisfied when one’s beliefs concerning what one has reason to do find an appropriate reflection in one’s desires” (Sayre-McCord 1997: 64). The two standards of rationality seem to be independent of one another (ibid., 65), for while “there does seem to be something irrational about failing to desire to do what one takes oneself to have reason to do . . . It can’t be because one is failing to desire as one has reason to desire” (ibid., 64). This logical independence of the two standards poses a problem for Smith because he claims that we can solve the moral problem (i.e., we can reconcile beliefs (1), (2), and (3)) by distinguishing between normative reasons for action and motivating reasons for action. Smith claims that his understanding of a normative reason is what will help to explain C2, and thus MI, while also providing us with a way to say that moral judgments express beliefs. But the only (or most) plausible interpretation of C2 involves the second standard of rationality, while it is the first standard of rationality that is involved in Smith’s understanding of a normative reason. So “[t]he worry is that Smith’s positive [Ideal Advisor] account of [normative] reasons seems not to contribute in any special way to an explanation of [why one ought to change one’s desires to fit one’s beliefs regarding what one ought to do]” (Sayre-McCord 1997: 65).

What follows are some of the details of Smith’s argument and some of the problems that Sayre-McCord identifies.

Smith claims that his account of normative reasons enables us to see why C2 is true. C2 says that if a person believes she has a reason to φ, then she rationally should desire to φ. The account of normative reasons says that a person has reason to φ in C if and only if were she fully rational she would want herself, as she actually is, to do it. So C2 becomes:

C2': If a person *believes that* if she were fully rational she would want herself, as she actually is, to φ in C, then she rationally should desire to φ.

The substitution into a belief context is justified, Smith says, because his account of normative reasons is an analysis. For Smith, an analysis of a concept is “a specification of the dispositions to act, judge, and infer, the possession of which constitutes competence with the concept in question” (Sayre-McCord 1997: 58 n. 2). If Smith’s analysis of normative reasons is true, the dispositions referred to by the left-hand side of the analysis are the very same dispositions referred to by the right-hand side of the analysis.

Why should we think that C2' is true? Why should we think that a person is irrational if they fail to be motivated to do what they believe to be morally required? What if their belief is false? If a person lacks the desire to do what she falsely believes she ought to do, wouldn’t it be “more rational” to correct the false belief rather than come to be motivated to act on that belief? After all, part of Smith’s conception of the fully rational agent is that she has no false beliefs. Of course, another part of that conception is that the agent have “a maximally coherent desiderative profile”.

Smith attempts to explain the rationality involved in C2' and C2 in terms of *coherence*, the idea being that an agent who fails to desire in accordance with her beliefs about what she ought to do suffers from a kind of “practical incoherence” (Sayre-McCord 1997: 70). He then tries to use his analysis of normative reasons to make clear the nature of this incoherence, claiming that the person who “believes she would desire herself to φ in C if she had a maximally coherent desiderative profile *but* has no desire to φ in C” is *less coherent* than the person who has the same belief and has the desire (ibid., 68). Once we accept this claim, Smith says, we understand why the person who believes she has a reason to φ in C but has no desire to φ in C is *less rational* than the person who believes that she has a reason to φ in C and has the desire to do so. But, as Sayre-McCord points out, it is not at all clear why one is more coherent as an agent if they desire in accordance with their beliefs about what they ought to do than if they should lack such desires. Again, what if the belief is false? what if the FPR agent would *not* desire that either agent φ in C? In that case isn’t it more plausible to say that the person who lacks the desire is more coherent (ibid., 73)?

Even if we have reasons for thinking that agents are always more coherent if they desire in accordance with their practical beliefs, why think that they are more practically *rational* in virtue of this coherence? What if the agent *disvalues* having a maximally coherent desiderative profile? For such an agent it would seem that the lack of the coherence in question would not be a mark against her practical rationality (ibid., 74). Sayre-McCord also offers another reason for thinking that there is a disconnect between being rational and having a maximally coherent and unified desiderative profile (ibid., 75-76): it is not at all clear that it makes sense to talk about coherence among desires because it is not at all clear that desires can justify and explain other desires. We might partially address this worry by talking instead about the evaluative beliefs associated with certain desires, and whether these beliefs contribute to the overall coherence of an agent’s beliefs. Coherence among desires could then be measured in terms of coherence among the associated evaluative beliefs (if there are any such beliefs associated with the desires). Even so, this still leaves us with a disconnect because “the pressure to have coherent desires is pressure provided by one’s evaluative beliefs—not some fetishistic concern that one’s desiderative profile be maximally coherent” (ibid., 76). Where is the pressure toward coherence among the rest of one’s desires, those desires which are not associated with any evaluative beliefs? Why is the pressure always in the direction of changing one’s desires rather than one’s beliefs?

Smith needs to explain the truth of C2, of why it is irrational to fail to desire to do what one believes one has a reason to do. He says that “the sort of irrationality involved is the irrationality of suffering a certain kind of incoherence” and that “his analysis (of what it is to believe one has a reason to φ in C) makes manifest, and so explains, just what the incoherence is” (Sayre-McCord 1997: 77). But Sayre-McCord argues persuasively that “there really is no incoherence involved in failing to desire what one believes one would desire if one had a maximally coherent desiderative profile”; he also argues persuasively that “if one were to insist on seeing the mismatch [between the desire and belief at issue] as a matter of incoherence, such incoherence as it would be does not plausibly reflect badly on one’s rationality” (ibid., 77). So Smith’s analysis of a normative reason neither explains the incoherence referred to nor the irrationality involved with C2.

Finally, Sayre-McCord points out that Smith’s account of the truth of C2 faces a major problem even if we were to grant that an agent suffers a kind of incoherence if she fails to desire to do what she believes she ought to do, and even if we were to grant that this agent is practically irrational in virtue of this incoherence. Even if Smith is right on these two counts, he “won’t have an explanation of the truth of C2 unless *everyone* who believes that they have reason to φ in C actually [i.e., consciously] believes that they would want themselves to φ in C if they had a maximally coherent desiderative profile” (ibid., 77). As Sayre-McCord observes, this follows because “On Smith’s view, a person who believes she has reason to φ in C and yet fails to desire to do it is irrational *because* she is believing she would want herself to φ in C if she had a maximally coherent desiderative profile” (ibid., 77). But this requirement also follows from Smith’s claim, made numerous times, that the agent “fails to be rational by her own lights” (Smith 1994: 62, 177, 178). Smith’s support for C2 is therefore highly problematic as it currently stands because it relies on an entirely implausible assumption. Smith does respond to this problem, but his response forces him to give up the claim that the agent fails to be rational by her own lights (1997: section VI). Smith’s response is such that it is no longer clear from where inside the agent the pressure will come from that leads her to adjust her desires to fit her beliefs about the normative reasons she has.4[[47]](#footnote-47)7 In the absence of such pressure, normative reasons will fail to have practical import, even in the FPR agent.

The above set of problems are just some of those facing Smith’s solution to the moral problem. More worries could be expressed: for example, about his understanding of moral facts as facts about the desires that FPR agents would converge on; or about his analysis of normative reasons. The problems we have looked at, though, are serious enough to make me think that Smith doesn’t have a solution to the moral problem. He hasn’t shown how the moral realist can subscribe to internalism, not even a very weak form of it. In his efforts to reconcile (1), (2), and (3), Smith gives up the core of (3), waters down (2) to the point of defending, arguably, a form of externalism (cf. Sayre-McCord 1997: 77, n. 26), and holds (1) in a way that is almost certainly incompatible with the realist’s mind-independence claim.

What lesson are we to draw from this? If Smith isn’t able to join realism with internalism even after giving up all of the ground that he does, it seems highly unlikely that a genuine form of realism can be made compatible with a genuine form of internalism. But if the moral realist must be an externalist, and if, as I have tried to show, IT is far more plausible than its opposite, then we have another reason to doubt the defensibility of moral realism.

*Conclusion to Chapters 6, 7, and 8*

At the very beginning of Chapter 6, and in earlier chapters, we noted that, whatever the problems facing moral realism, the thesis might still be defensible if the competing theories run up against even greater difficulties. In Chapter 6 I offered some reasons for thinking that there is a more viable alternative. The suggestion there was that we subscribe to a view of moral properties and moral judgments along the lines of what John McDowell has set out. If, as Smith and others have said, the central problem facing the realist and his opponents is that of making room, in their understanding of the nature of morality, for beliefs (1), (2), and (3), then it is worth seeing how the constructivist view I have argued for addresses this problem. If this view is a more viable alternative than realism, we would expect it to do a better job of making room for the three key beliefs. And I think it does.

Our constructivist will say that moral judgments are primarily cognitive in nature and do express the judge’s beliefs. But these judgments should also be seen as expressing noncognitive attitudes of the speaker, albeit in a very qualified sense. For moral judgments can be correct and incorrect. Yet they won’t be objective in the realist’s sense; they will be correct and incorrect relative to some evaluative framework. As for (2), our constructivist view makes room for it without qualification. When an agent judges that she ought to do X, we take it that she sincerely believes that she ought to do X. But this belief is evalutive in nature, and as such, it has a connection to motivation (necessarily so when the belief is sincerely held). Regarding (3), our constructivist view has room for the belief-desire theory of action and the Humean theory of motivation. To my mind, however, it is a mistake to say that all beliefs and desires (held with whatever strength) are distinct existences, for saying so precludes there being any necessary connection between them.

The realist cannot do as good a job incorporating these beliefs. True, he accepts (1) in an unqualified sense; but as we have seen, this is actually problematic in terms of explaining moral value experience, for in holding (1) in its unadulterated form, the realist cannot explain the other set of intuitions and beliefs that contribute to the tension described in the first pages of Chapter 1. The realist, as we have just seen in this chapter, has no room whatsoever for (2) if desires are understood as being crucial to motivation, and if the realist takes beliefs and desires to have different directions of fit. If the realist tries to make room for (2) by understanding moral judgments as having a noncognitive aspect to them, then the realist will have trouble holding onto (1). The best way for the realist to make room for (2) is by denying that desires are crucial to motivation. Of course, this forces the realist to give up on the core element of (3). It also forces the realist to be a rationalist, and rationalism seems quite implausible. Further, if we deny that desires are crucial to motivation, and yet take agents who are motivated to be goal-oriented, the realist has to explain how belief-states can have a direction of fit that we associate only with desires. Finally, if the realist makes room for (2), moral properties look to be queer entities. As for (3), the realist can allow that beliefs and desires are distinct existences, although this doesn’t seem to count in realism’s favor, as it seems to permit the existence of the reliable amoralist. And the realist has room for the belief-desire theory of action. But as we have seen, the realist has to give up the most important part of (3) if he is to hold (2). Since there are so many reasons for the realist to not hold (2), it seems that the realist would do best to hold (3) in its entirety, along with (1) in its entirety, and completely give up (2).

As I see it then, the constructivist makes room for (1), (2), and (3) exactly as we should, whereas the realist is really only capable of properly holding (3).

In my argument against externalism I noted that the internalist has no way of proving the existence of a necessary connection. She can only appeal to one’s intuitions regarding what we take sincere belief to involve, point to the problems facing the externalist, and show that internalism strongly coheres with the best, or most defensible, view of moral properties and moral judgments. My arguments against externalism never made much use of this third strategy. Yet I myself am also strongly inclined to accept IT exactly because of how well it seems to cohere with the other elements of my position, such as seeing moral judgments to be evaluative in nature. Much more, I think, could be said in this regard. For example, support for IT seems to come from conceiving of morality as aiming at human well-being. If we conceive of morality as aiming at human well-being, we will see our moral judgments as being goal-oriented—in the sense that they reflect or express our dispositions to promote human well-being. In other words, we will be inclined to see our moral judgments as having motivation already built into them.

Now, it is true that the realist can make a similar claim. Externalism strongly coheres with realism, so much so that the realist who is an externalist is likely to charge the internalist of begging the question against realism. While there is often some truth to this charge, the realist still needs to do more to make his case. For as we have seen in Chapters 6 and 7, the realist has problems explaining other common features of our moral beliefs and practices, and the realist who is an externalist still has to adequately explain, among other things, the phenomenology of moral language use. The realist needs to give us some other reason for being an externalist other than realism itself; otherwise the internalist can always legitimately charge the realist of begging the question. If the realist cannot make a separate case for externalism, then externalism rises and falls with realism. So any reasons we have for rejecting realism are reasons for rejecting externalism.

In any case, the reasons that I have offered in support of IT are, I think, mostly non-question-begging ones in the sense that they don’t presuppose what the realist would outright reject.4[[48]](#footnote-48)8 So if my arguments in support of IT are persuasive, we can legitimately say that we have an *additional* substantial reason for rejecting brands of moral realism committed to the two core tenets.

There is one final concern that we should perhaps return to. The realist and others will say that the constructivist position that I want to espouse is committed to “besires”. This was the “third problem” that was raised on page 280 of Chapter 6. It is a problem which any antirealist who subscribes to IT faces. The charge is that it is a mistake to understand a moral judgment like “I ought to do X” as a belief state which also has a necessary connection to a motivational state, where it is assumed that this motivational state involves the existence of a desire, or some desire-like state. If we say that the belief state and desire state are distinct existences, there is no reason to think that a necessary connection holds between them. But, it is said, if the two states are not distinct existences, then they must be a single, unified state—that is, a besire. And besires look to be quite mysterious: for how could a single state have two directions of fit that are exactly opposite of each other?

Proponents of IT who are antirealists have a way to avoid this commitment to besires. They can say that, while the utterance “I ought to do X” primarily expresses a belief state, when it is sincerely held it also expresses, or involves, a desire state. Our constructivist can say that beliefs and desires have opposite directions of fit, but deny that all beliefs and desires are distinct existences. We saw earlier that we cannot have certain desires without having certain beliefs; one doesn’t desire to go to law school unless one believes that there are law schools to go to. If this is true about desires, why can’t it also be true about beliefs? It is very implausible to think that beliefs exist in isolation from the other contents of our minds. One only needs to consider the role certain desires play in our coming to have certain beliefs. The likelihood of our coming to have certain beliefs is often a function of our having certain cares and interests; we are more reliable perceivers of injustice, for example, if we care about justice. Other things being equal, there is a much better chance that a house will be built correctly if it is done by someone who cares about doing it right. Then, too, we believe that beliefs influence the nature and shape of our dispositions. We also believe that beliefs can be held with varying degrees of strength, and that part of what determines the strength with which a belief is held is the degree to which the belief has left a mark on our dispositions. So we think it possible that our strongly held beliefs about what we ought to do are reflected in our dispositions to act. In fact, our sense of what is involved in the meaning of sincerely holding a moral belief says that this is not just possible, but necessary. The genealogy of certain strongly held beliefs involves certain desire states. This idea doesn’t seem mysterious at all.

With this third objection addressed, I think it is safe to say that there are at least three important features of morality that the moral realist cannot adequately account for—due to the fact that each of the features conflicts with the realist’s second core tenet. These features are important enough that we ought to conclude that any form of moral realism which holds the two core tenets is not a defensible thesis.

1. 1 This is true, anyway, if it turns out that the reasons for rejecting externalism are largely independent of the kinds of reasons for rejecting moral realism that were discussed in Chapters 6 and 7. In what follows, I try to argue for internalism without begging the question against the realist. [↑](#footnote-ref-1)
2. 2 One would have reason for holding this belief if it is the case that what makes something morally right or wrong is partly a function of those affective responses that are characteristic of us. In Chapter 6 I argued that this is in fact the case, but one of the reasons given was that we want to be capable of being motivated in the appropriate ways by what is morally right and wrong. If the realist has a way to explain the connection to motivation, or can argue that a necessary connection is not crucial, then there is less reason for thinking that those of our affective responses that are characteristic of us need to have a role in what makes something right or wrong. [↑](#footnote-ref-2)
3. 3 The judgment being that she ought to ϕ because ϕ-ing is morally required. [↑](#footnote-ref-3)
4. 4 It also aims to make clear that the crucial necessary connection is to motivating states. Generally speaking, internalism is a thesis about the extent to which our moral judgments have practical import. Part (i) of IT expresses a form of internalism about reasons and part (ii) expresses a form of internalism about motives. Part (ii) is by far the more controversial claim. [↑](#footnote-ref-4)
5. 5 Some internalists (Michael Smith, e.g.) hold that beliefs and desires are “distinct existences”. They claim that the belief expressed by “I ought to do X” or “Doing X is the morally right thing to do” can exist independently of the desire to do X, and vice versa. (Of course, some of these internalists would deny that the statements used in this example really express a belief.) This is a much stronger claim than one which says that we can distinguish between beliefs and desires in terms of “directions of fit” (see footnote 11 below). While IT is compatible with beliefs and desires having different “directions of fit”, the necessary connection to motivation that it involves does not exist if beliefs and desires are distinct existences. [↑](#footnote-ref-5)
6. 6 IT is obviously more difficult to defend if one is a cognitivist. While I have a cognitivist view of moral judgments, I also take them to express attitudinal states. My view is that beliefs and desires are not “distinct existences” even though we can properly think of them as having opposing “directions of fit”. Moral judgments should be seen as expressions of the speaker’s evaluative beliefs. [↑](#footnote-ref-6)
7. 7 Unless there is reason to explicitly use them, I will from here on drop the modifiers ‘consciously’ and ‘truly’ from ‘belief’. So in what follows if I say that a person believes something, I mean that they truly believe it, and consciously so. The need to talk about conscious belief becomes important when discussing Smith’s internalist claims. Smith (1997: 102-103) talks about agents believing that ‘if they were fully rational they would desire to do X’ even though they don’t consciously believe this; that is, these agents would not themselves describe the content of this belief state as “If I were fully rational I would desire to do X”.

   Incidentally, the claim that I make here is one that Brink denies (Brink 1986: 32). [↑](#footnote-ref-7)
8. 8 See, for example, Brink’s summary of a range of internalist positions in (Brink 1989: 40-41). [↑](#footnote-ref-8)
9. 9 In other words, in order for something to constitute a reason for someone, they have to see it in a certain light. But seeing that something in the necessary ‘certain light’ requires having certain attitudes and beliefs. [↑](#footnote-ref-9)
10. 10 For example, I think it is much harder to reject IT if we think that the best way to understand our moral judgments is along the lines that McDowell suggests, with the sort of entanglement described earlier. See pp. 257-59 of (Platts 1979) for how a realist might try to respond if the first necessary connection is granted. [↑](#footnote-ref-10)
11. 11 G. E. M. Anscombe employed this notion in her book *Intention* (Oxford UP 1963) to roughly distinguish between (in the words of Mark Platts) “two *kinds* of mental states, factual belief being the prime exemplar of one kind and desire a prime exemplar of the other . . . The distinction is in terms of the *direction of fit* of mental states with the world. Beliefs aim at the true, and their being true is their fitting the world; falsity is a decisive failing in a belief, and false beliefs should be discarded; beliefs should be changed to fit with the world, not vice versa. Desires aim at realisation, and their realisation is the world fitting with them; the fact that the indicative content of a desire is not realised in the world is not yet a failing *in the desire*, and not yet any reason to discard the desire; the world, crudely, should be changed to fit with our desires, not vice versa” (Platts 1979: 256-257).

    The reader may notice, in what follows, that my argument doesn’t rest on this second assumption. In mentioning it, I merely want to point out that I think IT is compatible with it. I would see it as a failing of IT if it weren’t compatible with this understanding of beliefs and desires. [↑](#footnote-ref-11)
12. 12 We can readily see why an antirealist view of an agent’s sincerely believing that she ought to ϕ lends itself favorably to IT whereas a realist interpretation does not. Antirealists often understand the judgment “I ought to ϕ” as something that already involves a motivating state. For example, many antirealists interpret this kind of judgment as one which expresses an attitude or desire of the agent. The necessary connection exists, then, almost as a matter of course. But the realist takes the moral judgment to be an expression of a belief state, something that is distinctively cognitive in nature. The presence of such a state doesn’t seem to entail motivation if we take motivation to crucially involve some noncognitive state like a desire, and if we take beliefs and desires to have opposite directions of fit. We will look at why the realist has trouble making room for IT in greater detail below. At the present moment I simply want to consider arguments in support of part (ii) of IT which don’t beg the question against the realist in such a direct fashion—that is, by taking an antirealist interpretation of moral judgments. IT is thought to be far more intuitive than its negation, so much so that it is a major reason why many moral philosophers reject realism. One would therefore expect that there are intuitions supporting IT which are not obviously antirealist. [↑](#footnote-ref-12)
13. 13 An example of a specific state: being motivated to see a movie this Friday night. An example of a nonspecific state: being motivated to go to graduate school. [↑](#footnote-ref-13)
14. 14 For proponents of IT, the concern is not about the agent being sincere in her expression of the belief, but about whether the agent sincerely believes. They particularly want to exclude cases of self-deception, where the agent thinks she believes she ought to ϕ, but doesn’t really believe it. A good substitute for ‘sincerely believes’, then, is ‘truly believes’; only this understanding of ‘sincere’ still doesn’t give us a clear sense of why it is that the adjective is not superfluous. Notice that the adjective is not superfluous when modifying ‘believes’ if we think it tells us something about the strength with which the agent holds the belief. [↑](#footnote-ref-14)
15. 15 Or it could be that I have the appropriate belief and desire, that both are a response to, or arise from, a very specific set of circumstances, but then the situation radically changes before I have a chance to act. Or it could be that present circumstances are unsuitable (or believed to be unsuitable), while the relevant belief and desire are general in nature and not tied to a specific, single set of circumstances. For example, a person might sincerely believe that she ought to go to graduate school, desire to go to graduate school, but fail to presently act in any direct sense on this belief and desire because (she believes that) conditions are not conducive for doing so. [↑](#footnote-ref-15)
16. 16 The unsympathetic philanthropist example aims to persuade us that there are times when agents are truly acting from a sense of duty rather than from nonmoral motives. Michael Smith believes that depressed, or despairing agents, are a counterexample to part (ii) of IT, but we will see in a moment why that is not the case. [↑](#footnote-ref-16)
17. 17 This is true even if the externalist has a way to identify the belief state in question without relying on our understanding of the concepts we presently use to refer to, or characterize, it. For the externalist has to be referring to a belief state that the internalist would describe as being sincerely held if the externalist is to hold a position that conflicts with that of the internalist. [↑](#footnote-ref-17)
18. 18 Brink (1986, 1989) points to the existence of sociopaths and amoralists to argue in support of externalism. Smith (1994) points to the existence of incontinent and depressed agents to argue against part (ii) of IT. McDowell (1979) responds to those who think the incontinent agent tells against a position like his which combines the internalism of IT with a cognitivist understanding of moral properties. [↑](#footnote-ref-18)
19. 19 If it is true that all of the reasons for not eating the cake have nothing to do with the agent’s desires and dispositions, the agent won’t see any of these reasons as reasons. She thus won’t believe that she ought to avoid eating the cake. [↑](#footnote-ref-19)
20. 20 What distinguishes the depressed agent from the incontinent agent is the reason for not ϕ-ing. The incontinent agent struggles with conflicting motivations; he desires to not ϕ or to do something incompatible with ϕ-ing. The depressed agent struggles with despair, something which may also have a physiological basis for it. [↑](#footnote-ref-20)
21. 21 Smith (1994: Chapter 3) sees David Brink as one such externalist. While Brink does say that there can be a reliable amoralist (1989: 59), he doesn’t argue against forms of internalism by trying to show the existence of the reliable amoralist. He simply says that the internalist is forced to explain the amoralist either by saying that the amoralist is using moral terms with inverted commas, or is misidentifying what morality actually requires of us (1989: 59). [↑](#footnote-ref-21)
22. 22 For an opposing view, see (Brink 1986: 32). [↑](#footnote-ref-22)
23. 23 Smith argues against the reliable amoralist in Chapter 3 of *The Moral Problem*. But that argument, and his other arguments against externalism, are severely flawed.

    Smith presents two separate arguments against externalism, but neither one is sound. Both rely on the mistaken premise that the externalist must say that a person cannot be committed to doing the right thing, read *de dicto*, and also be committed non-derivatively to honesty, justice, equality, the well-being of others, etc. (71-83). In arguing for this premise Smith claims that the externalist has to say that what explains a good and strong-willed person changing their moral motivations to accord with a change in their particular moral beliefs is the person’s being motivated to do the right thing, read *de dicto*, not *de re*. The externalist, Smith says, is committed to the (implausible) idea that being motivated to do the right thing is sufficient for a person being good. But the argument for this supporting premise only goes through if we ignore one of the two conceptions of rationality which Smith heavily relies on throughout his book. From the point of view of one conception of rationality he is right to say that “the mere fact that I have found reason to change my judgment gives me no reason to change this motive” (74). But if we are using the other conception of rationality that he works with, it would be irrational to not change the motive.

    Smith’s second argument against the externalist has two parts. The positive part of that second argument (pp. 85 ff.) is also flawed. For it relies on the assumption that it is rational to always desire to do what one believes one ought to do. See Sayre-McCord (1997) for why we want to reject this assumption. Ironically, this assumption—the idea that is at the heart of one of his two conceptions of rationality—directly conflicts with the idea that “the mere fact that I have found reason to change my judgment gives me no reason to change this motive.” [↑](#footnote-ref-23)
24. 24 Here we might also think in terms of what it takes in order to competently employ moral concepts. [↑](#footnote-ref-24)
25. 25 David Wiggins quotes this same passage from Hume in his “A Sensible Subjectivism” (1991a: 233). Wiggins observes that the phenomenon which Hume notes is not adequately accounted for by moral cognitivists. [↑](#footnote-ref-25)
26. 26 I.e., the idea that the normative force of moral requirements does not vary between individuals unless there are morally relevant differences between the individuals in question. [↑](#footnote-ref-26)
27. 27 But even in these other cases one would think that the reasoning is ultimately grounded in beliefs which have come into being through the perception of moral facts. [↑](#footnote-ref-27)
28. 28 See (Wallace 1990: 355, 367, 370). [↑](#footnote-ref-28)
29. 29 I am copying (1) - (3) verbatim from Smith’s book in order to do justice to his argument. See (Smith 1994: 12). Unless otherwise noted, when discussing Smith’s views all further page references are to his book, *The Moral Problem* (1994). [↑](#footnote-ref-29)
30. 30 For example, if the office tower in which employee X works is on fire, and X desires to escape the danger, why does X run up to the 50th floor rather than down to street level? This action would be explained if X also has the belief that at the 50th floor there is a skyway to the next building and that this is the quickest escape from danger. [↑](#footnote-ref-30)
31. 31 This will be the ‘is’ of identity, not constitution. Regarding this distinction, see Chapter 2. [↑](#footnote-ref-31)
32. 32 What follows are only some of the serious difficulties facing Smith’s project. [↑](#footnote-ref-32)
33. 33 Smith’s understanding of the FPR agent does seem to entail a much stronger version of internalism. By definition, FPR agents have no false beliefs, all relevant true beliefs, and they always deliberate correctly. So it would seem that if a person is FPR, they aren’t just motivated to do what they believe they ought to do; they actually do it. For it would seem to be a requirement of being FPR that one do what one truly ought to do. [↑](#footnote-ref-33)
34. 34 Smith himself may deny this. My point is that we should understand Smith’s presentation of (3) as involving three separate claims. This will enable us to salvage part of Smith’s project. As we will see shortly, there are two main problems for Smith that appear to be insurmountable: making room for (2); and holding (1) as the realist does. Smith’s arguments for MI don’t go through. [↑](#footnote-ref-34)
35. 35 As Wallace puts it, the principle “maintains that processes of thought which give rise to a desire (as ‘output’) can always be traced back to a further desire (as ‘input’), one which fixes the basic evaluative principles from which the rational explanation of motivation begins” (1990: 370). [↑](#footnote-ref-35)
36. 36 As R. Jay Wallace observes (1990: 373), Smith argues *against* rationalism (or certainly seems to be doing so) in his paper “The Humean Theory of Motivation” (*Mind* 1987). In *The Moral Problem* Smith argues in favor of a Humean theory of motivation and in favor of rationalism. In “The Definition of ‘Moral’” (1999: 42-43) Smith also presents an argument in support of rationalism. [↑](#footnote-ref-36)
37. 37 Smith argues as follows: To have a motivating reason is just to have a desire (along with some means-end beliefs) (1994: 130); for having a motivating reason involves having a goal, “having a goal is being in a state with which the world must fit,” and “being in a state with which the world must fit is desiring” (116). Desires, in other words, (along with means-end beliefs) constitute motivating reasons (125). By contrast, belief states cannot by themselves constitute a motivating reason, since belief states lack the necessary direction of fit (125). [↑](#footnote-ref-37)
38. 38 Although the view that beliefs and desires are distinct existences seems to conflict with Smith’s claim that the reliable amoralist is not possible. [↑](#footnote-ref-38)
39. 39 For example, he tells us that “moral facts are wholly determined by circumstances” (1994: 6, 11), that moral truths are non-relative and apply categorically, that moral properties are identical to natural properties, that by engaging with others in a process of systematic justification of our desires, we can end up with desires that in no way depend on the actual desires that we started with (1994: 173), and that these desires that we end up with are the moral facts that we seek (or they are the best evidence we can have of those moral facts (175)). When Smith characterizes his moral facts as non-relative, he implies that they are mind-independent, for Smith’s moral facts are not relative to the particular desires or mental states of any individual or any group of individuals. Smith is not advocating any sort of “morals by agreement” and he doesn’t see his view as constructivist. [↑](#footnote-ref-39)
40. 40 Besides, I don’t think it is a platitude that “what we have normative reason to do is what we would desire that we do if we were fully rational”. Nor do I think it a platitude that reasons for action are non-relative in the sense that Smith speaks of. [↑](#footnote-ref-40)
41. 41 See (Williams 1980) regarding internal reasons. This part of Smith’s position coheres well with his view that the reliable amoralist is not possible. [↑](#footnote-ref-41)
42. 42 The thought being that, if there can be different forms of life for human beings, one would think that there can be different normative frameworks regarding how we ought to live our lives. I am also inclined to believe that there can be different normative frameworks regarding how we ought to live our lives because I take it that there can be more than one set of political institutions under which human beings can flourish and because I take morality and politics to be strongly interconnected. [↑](#footnote-ref-42)
43. 43 In a review of Smith’s book, *Ethics and the A Priori* (2004), Simon Kirchin offers a couple of other important observations about the problems facing Smith in trying to combine non-relativism with internalism. (Many thanks to Norman Dahl for bringing this review to my attention.) It was actually Kirchin’s review that alerted me to this particular problem. Before seeing Kirchin’s review, I had always thought Smith’s argument for nonrelativism (1994: 166-174) was also supposed to be an argument *against* the claim that all practical reasons are internal ones. (I haven’t yet read Smith’s “Internal Reasons” paper.) [↑](#footnote-ref-43)
44. 44 As Sayre-McCord observes, “[Smith’s] view seems to fit poorly both with internalism and with the Humean account of practical reason. For, on the one hand, there appears to be no necessary connection of any sort between what we think we would desire were we fully rational (in Smith’s sense) and what we are actually motivated to do; while, on the other hand, a Humean would presumably resist the suggestion that the reasons we have for action might be grounded not in what we actually desire but in what we would desire were we, in striking ways, different than we actually are” (1997: 59). [↑](#footnote-ref-44)
45. 45 It is worth comparing this issue with what Wallace says is at stake between the Humean and the rationalist regarding the nature and scope of moral requirements (Wallace 1990: 356). The Humean regarding motivation, one who is also committed to an internalist moral psychology, has to say that the scope of moral requirements is dependent on agents’ existing desires or dispositions. The rationalist, on the other hand, holds that moral requirements hold categorically; the normative force of moral requirements doesn’t depend on agents’ antecedent desires. [↑](#footnote-ref-45)
46. 46 See (Smith 1994: 148). [↑](#footnote-ref-46)
47. 47 Regarding this pressure, see (Smith 1997: 86, 100, 119). Smith says that “Beliefs about normative reasons, when combined with an agent’s tendency to have a coherent psychology, can thus cause agents to have matching desires” (1997: 100). Why should we think this will be the case if, as Smith says, it is not the case that agents who consciously believe that they have a normative reason to do X consciously believe that if they were fully rational they would want themselves to do X? [↑](#footnote-ref-47)
48. 48 While I haven’t discussed them, Brink’s arguments in support of externalism fail to have any force at all against IT because they never even address it. When Brink argues against internalism, he argues against IT\* and its relatives (see (Brink 1986) and (Brink 1989: Chapter 3)). But one can espouse IT and still consistently reject IT\* and family. [↑](#footnote-ref-48)