Motivational Internalism and the Challenge of Amoralism[1]

Danielle Bromwich

Abstract: Motivational internalism is the thesis that captures the commonplace thought that moral judgements are necessarily motivationally efficacious. But this thesis appears to be in tension with another aspect of our ordinary moral experience. Proponents of the contrast thesis, motivational externalism, cite everyday examples of amoralism to demonstrate that it is conceptually possible to be completely unmoved by what seem to be sincere first-person moral judgements. This paper argues that the challenge of amoralism gives us no reason to reject or modify motivational internalism. Instead of attempting to diagnose the motivational failure of the amoral agent or restrict the internalist thesis in the face of these examples, I argue that we should critically examine the assumptions that underlie the challenge. Such an examination reveals that the examples smuggle in substantive assumptions that the internalist has no reason to accept. This argument has two important implications for the debate in moral motivation: first, it reveals that the motivational externalist needs a new argumentative strategy; and second, it shows that there is nothing especially problematic about a formulation of the thesis that captures the core internalist intuition that first-person moral judgements are necessarily accompanied by motivation.

1. Introduction

It is widely thought that moral judgements are necessarily motivationally efficacious.[2] This thought garners both popular and philosophical support from the phenomenology of moral experience. Imagine that today is Election Day and a friend of yours cannot decide whether to vote or not. Just as the polls are about to close, you convince your friend that he morally ought to vote, and so he forms the judgement 'I (morally) ought to vote today.' Suppose he then turns to you and says, 'Well, with that problem off my mind, I'm off home to bed!' Perhaps you say to him, 'Aren't you going to vote first?' Suppose that your friend replies, 'I know I (morally) ought to vote today, but I just don't see why that gives me any motive for doing so.' It seems fair to assume that this response would baffle you, and that is because we tend to think that the test of whether someone really judges an action to be right or to be done is whether that person is motivated to do it.[3]

Motivational internalism is the thesis that captures this commonplace thought about the practicality of moral judgements. It tells us that necessarily if I sincerely judge that 'I morally ought to φ' then I will be motivated to φ.[4] The moral judgement can be a truth-apt belief or a non-truth-apt conative state, but either way what is distinctive about the thesis is that motivation necessarily

accompanies the moral judgement.[5] But despite having its roots in commonsense morality, motivational internalism appears to be in tension with another aspect of our ordinary moral experience. Proponents of the contrast thesis, motivational externalism, cite everyday examples of amoralism to demonstrate that it is conceptually possible to sincerely judge that 'I morally ought to φ' and yet not be motivated to φ.[6] Psychopaths and depressives are often cited as real-life amoralists.[7] Such agents appear to be completely unmoved by what seem to be sincere first-person moral judgements. Let us call the apparent possibility of being unmoved by such judgements the challenge of amoralism.[8]

Motivational internalists tend to respond to this challenge in one of two ways. Some deny that amoral agents make moral judgements. R. M. Hare, for instance, famously claimed that amoralists just make 'inverted commas' judgements as opposed to sincere first-person moral judgements.[9] Others weaken their thesis to accommodate some examples of amoralism. Michael Smith, for instance, champions a *defeasible* conceptual connection between moral judgement and motivation in order to accommodate apathy and yet exclude other cases of amoralism.[10] Defeasible motivational internalism is the thesis that, necessarily if I sincerely judge that 'I morally ought to φ' *and* if all other things are equal, then I will be motivated to φ. Smith claims that 'depressions', which appear to leave one's ability to make sincere first-person moral judgements in tact but quash one's motivation, are 'a fact of ordinary moral experience'[11]; a fact that any plausible account of moral motivation should be able to accommodate. This defeasible formulation of the thesis allows Smith to claim that depressives make sincere moral judgements, but since things are not otherwise equal—that is, they are depressed—they are not motivated by them. Smith just denies that the same is true of amoralism. He claims that '. . . the *very best* we can say about amoralists is that they try to make moral judgements but fail.'[12]

Motivational externalists are less than impressed by these responses. They claim that not only do responses like Hare's fail to take their challenge seriously,[13] but responses like Smith's—which restrict internalism so that it can accommodate some but not all commonplace examples—reveal that it is the internalist who is under the burden to justify the assumption that one cannot have mastery of moral terms or concepts in the absence of motivation.[14] In fact, motivational externalists are so unimpressed that they have declared a victory of sorts in recent years: they claim to have shown that the burden of proof lies firmly on the shoulders of their opponent.[15]

This paper argues that victory has been declared prematurely because the challenge of amoralism gives us no reason to reject, modify or doubt motivational internalism. Attempting to diagnose the motivational failure of the amoral agent or restrict the internalist thesis in the face of these examples only succeeds in giving the challenge more credibility than it deserves. Instead we should critically examine the assumptions that underlie the challenge. When we do that we see that the examples of amoralism smuggle in substantive assumptions that the internalist has no reason to accept and can reject without assuming the truth of her own position.

The negative arguments that compose this critical examination have two positive implications for this debate in moral motivation. First, they imply that new versions of the challenge are unlikely to fare any better than those considered in this paper since what is required to disprove motivational internalism or at least shift the burden of proof is a successful defense of those smuggled assumptions. Second, they demonstrate that there is no reason to be enthusiastic about a restricted defense of motivational internalism. In recent years, the traditional nondefeasible formulation of the thesis has lost favor in light of the challenge of amoralism.[16] Removing this challenge removes the obstacle that stands in the way of accepting motivational internalism as it is traditionally understood.

2. The Challenge of Amoralism

The challenge of amoralism is designed to create a presumption in favor of motivational externalism. We are invited to consider familiar cases, free of any philosophical baggage, of agents who seem capable of making sincere first-person moral judgements without having the appropriate motivation. Consider the bully who fervently taunts an overweight classmate even though he ready accepts that he shouldn't do so. Or the lustful co-workers who eagerly embark on an affair even though they both truly believe that it is wrong to cheat on their partners. Or the mother with postpartum depression who is so overwhelmed with feelings of worthlessness that she cannot bring herself to keep to regular feedings even though she firmly believes that it is wrong to neglect her baby.

Proponents of the challenge are careful to construct their examples in a way that suggests that the featured agents are free of any cognitive impairments or failures of understanding that might be used to explain the lack of appropriate motivation.[17] Richard Joyce's version of the challenge provides a particularly nice case in point. His favored amoral agent is the purely evil agent; the agent who is motivated to act in ways that she must not (where this is interpreted *de re*).[18] He describes Eugenie, a Marquis de Sade character, who was once pure and innocent but becomes purely evil after being corrupted by a couple of sadistic libertines. Even so, Joyce emphasizes that

> Eugenie is, before her downfall, competent with moral predicates—indeed, she has been well brought up, and has a particularly sensitive moral sense. After her conversion at the hands of the diabolical Mme. de Saint-Ange, Eugenie applies those predicates as before: she calls acts of charity 'good,' acts of licentiousness 'wicked.' But her motivation has shifted: what she calls 'good' repels her and what she calls 'wicked' attracts her.[19]

Her competence with moral predicates implies that her cognition and moral understanding are unimpaired by her conversion experience; only her motivation has changed.

But these kinds of examples only constitute counterexamples to motivational internalism if two conditions are met. First, they must be plausible. If they are radically divorced from ordinary moral or psychological phenomena, it is often hard for us to conceive of the agent. Second, they must be incompatible with or at least problematic for internalism. If both parties can account for the phenomena in an equally plausible way, these examples have no normative force against motivational internalism.

Not all putative counterexamples to motivational internalism meet the first condition. It is commonly thought that the purely evil agent violates this condition.[20] After all, while it is plausible to think that one could be left cold by one's first-person moral judgement because one is temporarily indifferent to morality, it is not as plausible to think that one could be motivated to perform actions that are contrary to the prescription of one's first-person moral judgement because one is motivated by a desire that is derived from one's standing nonderived and noninstrumental desire to do evil. But many of them do. It is plausible to suppose that there could be agents who seem to be adept at making moral judgements but who are unmoved by these judgements. The real problem, however, is that the examples that meet the first condition fail to meet the second one. Close examination reveals that their plausibility lies in a feature that the motivational internalist can accept without difficulty.

An ambiguity lurks within all the apparent counterexamples. The examples are all of the same form: an agent, who seems to be adept at making moral judgements, makes what appears to be a genuine moral judgement but is unmoved by that judgement. But what constitutes being 'adept' at making a moral judgement? And what constitutes a 'genuine' moral judgement? The emphasis placed on the amoral agent's unimpaired understanding suggests an answer and constraint on any plausible example of amoralism: the agent has only made a genuine moral judgement if she understands the content of that judgement.[21]

However, motivational internalists and externalists have different conceptions of understanding. This can be seen by considering why the externalist claims that the amoralist understands the content of a first-person moral judgement, and would assent to it if asked. The evidence that the externalist adduces for this claim is that amoralists are just as competent as other speakers (or their past, morally better selves) in determining the extension of 'morally right'. That is, externalists appear to hold that if a subject can determine the extension of a moral concept contained within the content of her moral judgement correctly (for the most part), then that subject understands that judgement. But this is precisely what internalists deny. Internalists maintain that if a subject understands the content of her moral judgement, then that subject will be necessarily motivated by that judgement. But the purported counterexamples only have force if one accepts the motivational externalist's conception of understanding.

It should be clear why these examples are insufficient to constitute genuine counterexamples to motivational internalism. What is plausible about them is that a subject can determine the extension of moral concepts correctly (for the

most part) without being motivated by moral judgements that contain those concepts. The amoralist uses moral terms with the same extension as our moral terms, and properties that explain the amoralist's use of those terms are the very same properties that explain our use of the terms. But the internalist can accept this fact without difficulty. She just does not agree that this constitutes making a genuine moral judgement. Since the examples of amoralism are neither incompatible with nor problematic for motivational internalism they fail to constitute counterexamples to that thesis.

3. The Assumptions that Underlie the Challenges of Amoralism

One might think that the examples of amoralism still create a presumption in favor of motivational externalism. In what follows, I consider two ways they might do this, and argue that in each case they do not. I argue that what gives the examples the veneer of plausibility is two concealed but controversial assumptions that can be undermined without assuming the truth of motivational internalism. I close by considering and arguing against a challenge of amoralism that does not depend upon these assumptions, before drawing out two implications of my argument for this debate in moral motivation.

3.1. An Assumption about Understanding

I have argued that the examples of amoralism fail to constitute counterexamples because they depend upon an assumed conception of understanding; a conception that the motivational internalist does not share. But one might think that an argument in favor of this conception of understanding is unnecessary since we ordinarily think that if a subject can determine the extension of a concept correctly, then that subject is competent with the concept and, if all other things are equal, most likely understands the judgement that contains it. In fact, the examples of amoralism seem to succeed in revealing the plausibility of this assumed conception of understanding, which suggests that they do provide commonsense evidence against motivational internalism after all. However, since it is possible for an individual to be competent with a concept in just the way described without quite understanding it, the motivational internalist has reason to resist this assumed conception of understanding and deny that the examples do create the desired presumption.

In order to see this, consider the widely accepted account of communal meaning and norms for understanding provided by Tyler Burge in 'Intellectual Norms and the Foundations of Mind.'[22] Burge argues that the meaning of an empirically applicable term is fixed by social practice. For example, an ordinary term, like 'sofa', means the same as 'a piece of furniture [of such and such a construction] meant or made for sitting' because competent speakers, some of whom have reflected critically on the term, come to use the two expressions

interchangeably. While minimal linguistic competence simply consists in con-
forming to the practice of other competent speakers, greatest linguistic compe-
tence 'consists in abilities to draw distinctions, to produce precisifications, to
use numerous linguistic resources, to offer counterexamples to proposed
equivalences—that elicit the reflective agreement of other competent speakers.'[23]
What develops is 'a vast, ragged network of independence, established patterns
of deference which lead back to people who would elicit the assent of others.'[24]
Those who use and explicate ordinary terms well become persuasive to others
in their linguistic community, and over time their shared critical reflection on
archetypical applications of ordinary terms fix communal meaning of a term and
form the norms for understanding it. This shared dialectic provides the linguistic
community with 'normative characterizations' of ordinary terms. These are
statements about what is necessarily or essentially true about the object to which
the term applies. It is through this kind of social practice that 'a piece of furniture
[of such and such a construction] meant or made for sitting' comes to be a
normative characterization and therefore a synonym for 'sofa'.

   Burge's analysis has important implications for our debate in meta-ethics: it
reveals the possibility of a speaker reliably determining the extension of a
concept without understanding it. Burge asks us to consider *A*, a member of our
linguistic community who has developed a mastery of the English language.
A—let's call him Albert, for ease—has picked up the usual platitudes about
'sofas' and he can use the term reliably. However, after some reflection, Albert
develops doubts about sofa platitudes. He comes to the conclusion that sofas are
not items of furniture, rather pieces of artwork or religious artifacts.

> He believes that the usual remarks about the function of sofas conceal,
> or represents a delusion about, an entirely different practice. [Albert]
> admits that some sofas have been sat upon, but thinks that most sofas
> would collapse under any considerable weight and denies that sitting is
> what sofas are pre-eminently *for*. [Albert] may attack the veridicality of
> many of our memories of sofas being sat upon, on the grounds that the
> memories are products of the delusion.[25]

   In other words, Albert develops a nonstandard theory about what sofas are,
which is designed to challenge the community's normative characterization of
the concept. While this nonstandard theory does not impair his ability to
determine the extension of the concept in question correctly, it does lead Albert
to deny what is essentially or necessarily true about 'sofas'—that is, that 'sofas'
are items of furniture, which are designed to be sat upon.[26] By denying what *is*
essentially or necessarily true about the object to which the ordinary term
applies, Albert fails to understand the concept of sofa. After all, as Burge argues,
the linguistic community's normative characterization of a concept 'provide[s]
linguistic meaning—set[s] a norm for conventional linguistic understanding.'[27]
We must also suppose that Albert's attempts to persuade other competent
speakers of his normative characterization of the concept in question would fail,
and that he would stand corrected.[28] Other competent speakers in his community

would surely take him to be confused, mistaken, in the grips of (albeit) a sophisticated delusion about the concept.

One might object that Albert's unusual beliefs about sofas would influence his view of the extension of the concept. For example, on seeing a person sitting on a sofa-like object, Albert might deny that *that* object is a 'sofa'. However, since Albert readily admits that some sofas have been sat upon, he is more likely to dismiss these kinds of cases as anomalies. He may even argue that his theory can explain such cases. For example, he may think that some especially mischievous conceptual artists design their sofas to hold a considerable amount of weight. This accounts for how these pieces of artwork occasionally stand up to being sat upon. Since Albert is aware that he is challenging the conventional meaning or normative characterization of the concept of sofa, it is likely that he will attempt to explain how (what he considers) the usual delusions about sofas have arisen and why they are mistaken.[29] He is therefore likely to assume that cases in which sofas have been sat upon neither falsify his theory nor require him to refine the concept in a way that would influence his extensional use of it. The *possibility* of an agent like Albert, who can reliably determine the extension of a term like 'sofa' without possessing the correct normative characterization of the concept, should lead us to doubt the externalist's assumption that reliably determining the extension of a concept correctly is a sufficient condition for understanding that concept.

Burge helpfully points out that:

> Nearly anything can be the topic of non-standard theorising. Similar thought experiments apply to knives, clothing, rope, pottery, wheels, boats, tables, watches, houses. Both technical and everyday natural-kind notions clearly fall within the domain of the argument . . . Concepts of other ordinary objects and stuffs, which are not natural kinds are equally good examples: earth, air, fire, mountains, rivers, bread, food, dung. Notions associated with common verbs are also subject to strange theory.[30]

The broad scope here is significant: while Burge does not use a moral concept in his example, it is clear that the thought experiment is applicable to moral concepts. This example reveals that it is a general feature of understanding—not a special feature of *moral* understanding—that one can reliably determine the extension of a concept correctly without understanding that concept. This example, then, undermines the externalist's conception of understanding without engaging in any special pleading because it does not depend upon anything particular being true of morality nor does it depend upon the truth of motivational internalism.[31]

It is worth noting that there appears to be a salient difference between Albert and the amoralist. That is, while it is clear that Albert has developed a nonstandard theory of the concept 'sofa' which explains his lack of understanding, it is not clear that the amoral agent has—or necessarily has—a nonstandard theory of moral concepts.

But since Burge's example is not being used to explain the amoralist's moral motivational inertia or to make the positive claim that all amoralists are nonstandard theorizers, this point does not undermine the argument. Burge's example is just being used to undermine the externalist's assumed conception of understanding by showing that determining the extension of a concept and understanding that concept can come apart. This suggests that it is possible to determine the extension of a concept correctly without understanding that concept. And since Burge's example does not depend upon the truth of motivational internalism, the internalist can conclude—without begging the question—that the externalist's assumed conception of understanding is not plausible enough to create the desired presumption.

What the Burge example shows is that motivational externalist's conception of understanding is controversial, and that deflects a way of using the challenge of amoralism to create a presumption in favor of motivational externalism. And it does so without assuming the truth of motivational internalism.

3.2. An Assumption about Motivation

There is another difference between Albert and the amoralist that might create a presumption in favor of motivational externalism. Albert lacks certain linguistic dispositions. It is likely that he is disposed to think and say odd things about sofas and is disposed to think that standard things said about sofas are odd. The amoralist, on the other hand, might be disposed to think and say perfectly standard and sensible things about morality and disposed to think that the standard things said about morality are perfectly sensible. The only thing that appears to be missing from the amoralist's repertoire is certain behavioral dispositions. That is, dispositions to act in accordance with what appear to be sincere first-person moral judgements. But why think that lacking behavioral dispositions amounts to a failure to make a genuine moral judgement?

There seems to be good reason to think that the internalist will have difficulty providing an answer to this question that is free of internalist assumptions. As Russ Shafer-Landau points out, the internalist will want to insist that 'the meaning of a moral judgement is given primarily or exclusively by an intrinsically motivating attitude.'[32] But insisting that the amoralist has not made a moral judgement *because* 'moral judgement' means 'a judgement that is necessarily accompanied by motivation' just involves assuming the truth of motivational internalism.

However, since the worry itself goes beyond the scope of motivational internalism, the internalist can appeal to independent meta-ethical commitments to provide a nonquestion begging response. In order to see this, notice that despite being a thesis about the motivational nature of moral judgements, motivational internalism is actually silent on a number of issues concerning the nature of these judgements. It just tells us that there is a necessary connection between sincere, first-person moral judgements and motivation. It does not tell

us whether moral judgements are cognitive or noncognitive mental states. It also does not tell us whether moral judgements necessarily motivate because they are partly constituted by motivational states or because they are necessarily accompanied by such states. But it is these additional meta-ethical commitments—that is, commitments that are not entailed by the truth of motivational internalism— that shape internalists' responses to this kind of worry.

Let me illustrate this with two examples. First, many motivational internalists are also noncognitivists.[33] They maintain that moral judgements are noncognitive mental states, and such mental states are intrinsically motivating. It is therefore the noncognitive nature of the moral judgement that explains why the amoralist's lack of appropriate motivation is evidence that he has not made a genuine moral judgement. Since it does not follow from the truth of motivational internalism that a moral judgement is a noncognitive state, this response to the worry does not beg the question.

Second, some motivational internalists are cognitivists.[34] They maintain that moral judgements are cognitive mental states. But since it is widely accepted that cognitive states are motivationally inert, the worry seems to be more problematic for internalists of this sort. The amoralist lacks the very behavioral dispositions we would expect him to lack if the moral judgement was a cognitive mental state. But if this is the worry, it is not a worry for motivational internalism. After all, when cognitivist motivational internalists respond to this kind of challenge, they appeal to meta-ethical commitments that are independent of motivational internalism. For example, some argue that cognitivist mental states are intrinsically motivating.[35] This involves arguing that the Humean theory of motivation—the thesis that cognitive states cannot motivate without the assistance of a conceptually independent desire[36]—is false.[37] Such a response does not beg the question against the motivational externalist. After all, the claim is that the cognitive nature of the moral judgement explains why the amoralist's lack of appropriate motivation is evidence that he has not made a genuine moral judgement. Since it does not follow from the truth of motivational internalism that a moral judgement is a cognitive state of this sort, this response does not beg the question.

The argument in this section is not designed to vindicate any particular conception of motivational internalism. It is just designed to undermine a motivational assumption that underlies a particular use of the challenge of amoralism. The assumption is that the motivational internalist cannot explain why lacking behavioral dispositions amounts to a failure to make a genuine moral judgement without thereby assuming the truth of her own position. But since motivational internalism is impressively silent on a number of issues concerning the nature of moral judgements, the internalist must appeal to other meta-ethical commitments to explain why the lack of motivation indicates that the amoralist has not made a sincere moral judgement. This reveals that the motivational internalist is not forced to beg the question in responding to this use of the amoralist challenge; she can help herself to independent meta-ethical commitments that are not entailed by the truth of motivational internalism to construct a reply.

3.3. The Third Refined Challenge of Amoralism: Svavarsdottir's Argument

At least one motivational externalist will agree that the forgoing uses of the challenge neither disprove motivational internalism nor shift the burden of proof.[38] Sigrun Svavarsdottir argues that this is because the debate has not shifted from a philosophical investigation about the nature of moral judgements to an empirical investigation of observable behavior. When it does, she claims that it becomes clear that the burden of proof lies with the motivational internalist.

The phenomenon she selects to be explained is that of a moral cynic called Patrick.[39]

> *Example of Patrick:* Virginia has put her social position at risk to help a politically persecuted stranger because she thinks it is the right thing to do. Later she meets Patrick, who could, without any apparent risk to himself, similarly help a politically persecuted stranger, but who has made no attempt to do so. Our morally committed heroine confronts Patrick, appealing first to his compassion for the victims. Patrick rather wearily tells her that he has no inclination to concern himself with the plight of strangers. Virginia then appeals to explicit moral considerations: in this case, helping strangers is his moral obligation and a matter of fighting enormous injustice. Patrick readily declares that he agrees with this moral assessment, but nevertheless cannot be bothered to help. Virginia presses him further, arguing that the effort required is minimal and, given his position, will cost him close to nothing. Patrick responds that the cost is not really the issue, he just does not care to concern himself with such matters. Later he shows absolutely no sign of regret for either his remarks or failure to help.[40]

Both parties can explain Patrick's behavior. The motivational internalist claims that if Patrick makes a sincere moral judgement, he is motivated to act in accordance with the judgement, despite his claims to the contrary. Patrick's motivation does not result in any observable moral behavior because either the judgement is not genuine or the appropriate motivation is defeated by a competing disposition to do otherwise. The motivational externalist claims that Patrick makes a sincere moral judgement but he is completely unmoved because he has no conceptually independent desire to act morally.

While it might appear as if we have reached a stalemate once again, Svavarsdóttir is quite convinced that the burden of proof has been shifted. She claims that the internalist is in trouble because her explanation of Patrick's behavior rules out the externalist's explanation. After all, the internalist cannot accept the possibility that Patrick has made a sincere moral judgement without having some motivation, as the externalist suggests. The internalist must hold that the externalist's hypothesis 'cannot be the right conclusion to draw under any circumstances',[41] which means that she must deny that this competing

hypothesis is even in the running. But if Svavarsdóttir is right about this, then the internalist seems to have violated a general methodological principle that governs empirical investigations. The principle tells us that an explanatory hypothesis incurs a burden of proof if it excludes a competing explanatory hypothesis that is not obviously unreasonable or defective. One cannot simply rule out competing but otherwise reasonable explanations just because they do not fit with our own when explaining some observable phenomenon. But in virtue of violating this principle, Svavarsdóttir concludes that the burden of argumentation shifts on to the internalist.

But this challenge of amoralism also fails. Svavarsdóttir is only able to shift the burden of proof in virtue of the observable phenomenon she selects to be explained. If we change the observable phenomenon from an agent who *is not* motivated in accordance with his moral judgement to an agent who *is* motivated in accordance with her moral judgement, we see that the burden of proof shifts on to the externalist.

Consider a modified version of Svavarsdóttir's original example.

> *Example of Patricia:* Virginia risks her social position, once again, to help a politically persecuted stranger because she thinks it is the right thing to do. Later she meets Patrick's twin sister, Patricia, who, like Patrick, could help a politically persecuted stranger without any risk—social or otherwise—to herself. Virginia, once again, starts her plea by first appealing to Patricia's compassion for the victims. Patricia confesses that she was unaware of the plight of these particular strangers. Virginia then appeals to explicit moral considerations: she tells Patricia that, since she now is aware of the plight of these politically persecuted strangers, it is her moral obligation to help these strangers; fighting their persecution is a matter of fighting enormous injustice. Patricia, like her brother, readily declares that she agrees with this moral assessment, but, unlike her brother, she is motivated to help these strangers. Later Patricia joins Virginia in helping these politically persecuted strangers.

Once again both parties can explain Patricia's behavior, but this time it is the externalist's hypothesis that rules out the internalist's hypothesis. While the externalist must deny that Patricia is necessarily motivated by a sincere moral judgement, the internalist can accept that Patricia is motivated in accordance with her moral judgement *either* because she has made a sincere moral judgement *or* because she has made a moral judgement and has a conceptually independent desire to be moral. Otherwise put, the internalist can accept either her own or the externalist's explanation as a coherent hypothesis whereas the externalist cannot accept the internalist's explanation as a coherent hypothesis because that very hypothesis is ruled out by her motivational thesis. In sum, Svavarsdóttir ultimately fails to shift the burden of argument. The result she obtains is not a general one; it is particular to the example she selects.

---

4. Implications for the Debate in Moral Motivation

The upshot of this critical examination of the challenge of amoralism is not entirely negative. The forgoing arguments have two important implications for the debate about moral motivation in meta-ethics.

First, the arguments suggest that motivational externalists need a new argumentative strategy. Arguments for motivational externalism are rarely positive.[42] Instead of extolling the virtues of their own account, externalists tend to rely on the challenge of amoralism to show that their position is at least free of the problems that plague the alternative. And this strategy has only become more popular with the rise of experimental philosophy.[43] But the arguments in this paper imply that the challenge of amoralism is not going to yield the desired result all by itself. It will continue to be either insufficient to create a presumption in favor of motivational externalism or unnecessary if the assumptions that underlie these examples are successfully defended. Perhaps the argument in this paper gives externalists the nudge that they need to provide a positive defense of their position; a defense that gives us reason to get excited about their account of moral motivation.

Second, the arguments have an interesting implication for motivational internalists as well. Motivational internalism is traditionally formulated to capture the following necessary and *nondefeasible* conceptual connection between moral judgement and motivation: necessarily, if I sincerely judge that 'I morally ought to φ', then I will be motivated to φ. This formulation has its roots in commonsense morality, in that we are pre-theoretically disposed to think that moral judgements are states that motivate those who make them sincerely. But the challenge to account for everyday examples of amoralism has led those sympathetic to internalism to abandon this traditional nondefeasible formulation in favor of the following defeasible one: necessarily if I sincerely judge that 'I morally ought to φ' and if all other things are equal, then I will be motivated to φ. As Michael Smith says

> '. . . though there is a conceptual connection between moral judgement and the will, the connection involved is the following defeasible one. 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. In other words, agents who judge it right to act in various ways are so motivated, and necessarily so, absent the distorting influences of weakness of the will and other similar forms of practical unreason on their motivations'[44]

Only a defeasible formulation has the necessary gap between the moral judgement and motivation that allows the internalist to take the examples of 'weakness of the will' into account.

But restricting the internalist thesis in the face of the challenge of amoralism—by maintaining that the conceptual connection between moral judgement and motivation is just a defeasible one—is quite a concession to the motivational externalist.[45] Motivational internalism is supposed to capture the

intuition that moral judgements are the kinds of states that necessarily motivate those who make them sincerely. But once the *ceteris paribus* clause is inserted between moral judgement and motivation, it is no longer clear that this is the case; it is now possible to make a sincere, first-person moral judgement and not be motivated. It becomes very difficult to distinguish this weakened internalist thesis from motivational externalism. After all, motivational externalism is the view that sincere moral judgements are not necessarily accompanied by motivation.[46] What the forgoing critical examination of the challenge reveals is that the motivational internalist has no reason to restrict her thesis to accommodate these apparent cases of amoralism. The arguments in this paper clear away an obstacle that has been thought to stand in the way of accepting the traditional necessary and nondefeasible formulation of the internalist thesis.

5. Conclusion

The arguments in this paper offer the motivational internalist a simpler line of defense against the ubiquitous challenge of amoralism. Instead of attempting to diagnose the motivational failure of the amoral agent or weaken the thesis in the face of these challenges, the internalist is urged to question the assumptions that underlie the purported counterexamples. This does not involve denying that the examples of amoralism are intuitive or arguing that they are but only as a result of an inference from a question-begging principle. It just involves a critical examination that reveals that the examples' veneer of plausibility resides in two concealed but controversial assumptions about understanding and motivation. These assumptions can be denied without assuming the truth of motivational internalism. If this is right, then the challenge of amoralism does not succeed in creating a presumption in favor of motivational externalism, and it does not give the motivational internalist any reason to doubt or modify her thesis. Moreover, stripping the normative force from these challenges has two important implications for the debate in moral motivation. First, it reveals that the motivational externalist needs a new argumentative strategy; and second, it shows that there is nothing especially problematic about a nondefeasible formulation of the thesis that captures the core internalist intuition that sincere, first-person moral judgements are necessarily accompanied by motivation.

Danielle Bromwich
Department of Philosophy
University of Massachusetts, Boston
Massachusetts, USA
daniellebromwich@gmail.com

NOTES

Gurpreet Rattan, Jonathan Weisberg, Tom Hurka and an anonymous referee for this journal. Earlier drafts of this paper were presented to audiences at the Department of Philosophy at the University of Toronto in December 2009, the Department of Philosophy at the University of Northern Illinois in January 2010, the Department of Philosophy at Fordham University in February 2010, the Pacific Division APA in April 2010, the New Scholars in Bioethics in May 2011 and the Department of Philosophy at Dalhousie University in June 2011. I thank the audiences at all these venues for their feedback.

² C. L. Stevenson, 'The Emotive Meaning of Ethical Terms', *Mind* 46 (1937), pp. 14–31; W. D. Falk, ''Ought' and Motivation', *Proceedings of the Aristotelian Society* 48 (1947–48), 111–38; R. M. Hare, *The Language of Morals* (Oxford: Oxford University of Press, 1952); T. Nagel, *The Possibility of Altruism* (Princeton, NJ: Princeton University of Press, 1970); J. Mackie, *Ethics: Inventing Right and Wrong* (New York: Penguin, 1977); J. McDowell, 'Virtue and Reason', *Monist* 62 (1979), pp. 331–50; M. Platts, *Ways of Meaning* (London: Routledge and Kegan Paul, 1979); S. Blackburn, *Spreading the Word* (New York: Oxford University Press, 1984); D. McNaughton, *Moral Vision* (Oxford: Basil Blackwell, 1988); D. Wiggins, 'Moral Cognitivism, Moral Relativism and Motivating Moral Beliefs', *Proceedings of the Aristotelian Society* 91 (1990–1991), pp. 61–85; J. Dancy, *Moral Reasons* (Oxford: Basil Blackwell, 1993); M. Smith, *The Moral Problem* (Oxford: Blackwell, 1994); M. Little, 'Virtue as Knowledge: Objections from the Philosophy of Mind', *Nous* 31 (1997), pp. 59–79; T. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998); M. van Roojen, 'Humean and Anti-Humean Internalism about Moral Judgements', *Philosophy and Phenomenological Research* 65 (2002), pp. 26–49; D. Bromwich, 'Clearing Conceptual Space for Cognitivist Motivational Internalism', *Philosophical Studies* 148 (2010), pp. 343–67.

³ These kinds of examples provide intuitive support for motivational internalism, but they are far from decisive. See S. Finlay, 'Value and Implicature', *Philosophers' Imprint* 5 (2005), pp. 1–20; and C. Strandberg, 'The Pragmatics of Moral Motivation', *Journal of Ethics* 15 (2011), pp. 341–69. The argument in this paper is not designed to demonstrate that they are decisive; it is just designed to show that the challenge of amoralism is not.

⁴ This is a rough formulation of the thesis. For more on various formulations and version of motivational internalism, see S. Darwall, *Impartial Reason* (Ithaca, NY: Cornell University Press, 1983), pp. 51–5; S. Darwall, 'Internalism and Agency', *Philosophical Perspectives* 6 (1992), pp. 155–74; C. Miller, 'Motivational Internalism', *Philosophical Studies* 139 (2008), pp. 234–36; and F. Bjorklund, G. Bjornsson, J. Eriksson, R. Francen Olinder & C. Strandberg, 'Recent Work on Motivational Internalism', *Analysis* 72 (2012) pp. 124–37.

⁵ Some maintain that motivation arises solely from the moral judgement itself. On this view, moral judgements are at least partially constituted by motivational attitudes. As Frankena puts it, 'The question is whether motivation is somehow to be 'built into' judgements of moral obligation, not whether it is to be taken care of in some way or other', p. 41. W. Frankena, 'Obligation and Motivation in Recent Moral Philosophy', in *Essays in Moral Philosophy*, ed. A.I. Melden (Seattle: University of Washington Press, 1958), pp. 40–81. Also, see D. Brink, 'Moral Motivation', *Ethics* 108 (1997), pp. 7–8. But motivational internalism just requires a necessary connection between moral judgement and motivation; it does not specify the nature of that necessary connection. It is therefore compatible with the view that moral judgements are not constituted by motivational attitudes but are necessarily accompanied by them. See J. Tresan, 'De Dicto Internalist Cognitivism', *Nous* 40 (2006), pp. 143–65; C. Strandberg, 'The Pragmatics of Moral Motivation', *Journal of Ethics* 15 (2011), pp. 341–69; and F. Bjorklund, G. Bjornsson, J. Eriksson, R. Francen Olinder & C. Strandberg, 'Recent Work on Motivational Internalism', *Analysis* 72 (2012) pp. 124–37.

⁶ The everyday examples include cases of amoralism and apathy. See, for example, M. Stocker, 'Desiring the Bad: An Essay in Moral Psychology', *The Journal of Philosophy* 76 (1979), pp. 738–53; D. Brink, *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press, 1989). p. 46; M. Smith, *The Moral Problem* (Oxford: Blackwell, 1994); A. Mele, 'Internalist Moral Cognitivism and Listlessness', *Ethics* 106 (1996), pp. 727–753; S. Svavarsdóttir, 'Moral Cognitivism and Motivation', *The Philosophical Review* 108 (1999), pp. 161–219; R. Shafer-Landau, 'A Defense of Motivational External-ism', *Philosophical Studies* 97 (2000), pp. 267–91; R. Joyce *The Myth of Morality* (Cambridge: Cambridge University Press, 2001); A. Roskies, 'Are Ethical Judgements Intrinsically Motivational? Lessons from Acquired Sociopathy', *Philosophical Psychology* 16 (2003), pp. 51–66; J. Doris and S. Stich, 'As a Matter of Fact: Empirical Perspectives on Ethics', in *The Oxford Handbook of Contemporary Ethics*, eds. F. Jackson and M. Smith (Oxford: Oxford University Press, 2005); C. Miller, 'Motivational Internalism', *Philosophical Studies* 139 (2008), pp. 233–55.

⁷ See M. Stocker, 'Desiring the Bad: An Essay in Moral Psychology', *The Journal of Philosophy* 76 (1979), pp. 738–53; D. Brink, *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press, 1989); M. Smith, *The Moral Problem* (Oxford: Blackwell, 1994); A. Roskies, 'Are Ethical Judgements Intrinsically Motivational? Lessons from Acquired Sociopathy', *Philosophical Psychology* 16 (2003), pp. 51–66; J. Doris and S. Stich, 'As a Matter of Fact: Empirical Perspectives on Ethics', in *The Oxford Handbook of Contemporary Ethics*, eds. F. Jackson and M. Smith (Oxford: Oxford University Press, 2005).

⁸ I use the term 'amoralism' broadly enough to capture all cases in which an individual makes a seemingly sincerely first-person moral judgement but fails to be appropriately motivated.

⁹ R. M. Hare, *The Language of Morals* (Oxford: Oxford University of Press, 1952), especially 124–6 and 163–5.

¹⁰ Smith claims that '. . . though there is a conceptual connection between moral judgement and the will, the connection involved is the following defeasible one. 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. In other words, agents who judge it right to act in various ways are so motivated, and necessarily so, absent the distorting influences of weakness of the will and other similar forms of practical unreason on their motivations' p. 61. M. Smith, *The Moral Problem* (Oxford: Blackwell, 1994). Dreier also cautions against overstating internalism. See J. Dreier, 'Internalism and Speaker Relativism', *Ethics* 101 (1990) pp. 6–26.

¹¹ Smith, *The Moral Problem* (Oxford: Blackwell, 1994), p. 120.

¹² Smith, *The Moral Problem* (Oxford: Blackwell, 1994), p. 68.

¹³ See D. Brink, *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press, 1989) and R. Joyce, *The Myth of Morality* (Cambridge: Cambridge University Press, 2001), p. 21.

¹⁴ Critics deny that Smith succeeds in providing a sound defense of this claim. See, for example, A. Miller, 'An Objection to Smith's Argument for Internalism', *Analysis* 56 (1996), pp. 169–74; H. Lillehammer, 'Smith on Moral Fetishism', *Analysis* 57 (1997), pp. 187–95; Svavarsdóttir, 'Moral Cognitivism and Motivation', *The Philosophical Review* 108 (1999), pp. 161–219.

¹⁵ In particular, see S. Svavarsdóttir, 'Moral Cognitivism and Motivation', *The Philosophical Review* 108 (1999), pp. 161–219 and R. Shafer-Landau, 'A Defense of Motivational Externalism', *Philosophical Studies* 97 (2000), pp. 267–91.

[16] See J. Dreier, 'Internalism and Speaker Relativism', *Ethics* 101 (1990), pp. 6–26; M. Smith, *The Moral Problem* (Oxford: Blackwell, 1994); C. Miller, 'Motivational Internalism', *Philosophical Studies* 139 (2008), pp. 233–55.

[17] In particular, see M. Smith, *The Moral Problem*, Oxford: Blackwell (1994), p. 123, and S. Svavarsdóttir, 'Moral Cognitivism and Motivation', *The Philosophical Review* 108 (1999) pp. 177–78.

[18] R. Joyce, *The Myth of Morality* (Cambridge: Cambridge University Press, 2001), pp. 21–3. For a similar challenge, see J. Dreier, 'Internalism and Speaker Relativism', *Ethics* 101 (1990), pp. 10–11.

[19] R. Joyce, *The Myth of Morality* (Cambridge: Cambridge University Press 2001), p. 21 Joyce adds, 'Eugenie knows . . . that she was raised to be particularly morally sensitive, and knows herself to be a more reliable judge than those around her, so even if everybody else in the world were to judge that some act of hedonism is permissible, if Eugenie judges herself to know better, then she would still want to perform that act', pp. 21–2.

[20] Joyce takes care in selecting his evil agent because examples of evil are vulnerable to two kinds of internalist explanations. First, the internalist can argue that the evil agent is actually akratic. Second, she can argue that the evil agent performs an evil action so as to attain some further end, 'and therefore is not doing evil for evil's sake.' R. Joyce, The Myth of Morality (Cambridge: Cambridge University Press, 2001), p. 20. So Joyce advances an example that is not vulnerable to either internalist rejoinders. He claims:

> . . . there's evil and there's *evil.* About a certain kind of agent—which for convenience I'll call 'purely evil'—one must be a success theorist. This agent is defined entirely in terms of her intentions and motivations. But 'pure evil' is a term of art. The kind of evil agent we more familiarly speak of not only has bad intentions, but acts, or intends to act in ways that *she must not* (where this is interpreted as *de re.*) A full-blooded *moral* assertion of 'S is evil,' in other words, holds (i) there are things that S must not do, and (ii) S intentionally does (or at least is motivated to do) these things. We might add for certain agents: (iii) S does then because S judges them forbidden. The purely evil agent, by contrast, is under a misapprehension: she believes that there are things that she must do, and she is motivated to do them, but her belief is mistaken. The error theorist need not deny the existence of the purely evil agent, but does deny the 'ordinary' evil agent captured by (i)–(iii). (p. 21)

But this conception of pure evil is normally reserved for devils and not for human beings. Even Drier, who advances a similar challenge, acknowledges that these agents are not commonplace. J. Dreier, 'Internalism and Speaker Relativism', Ethics 101 (1990), p. 11. Ronald Milo, in his analysis of the concept of immorality, says:

> . . . the most extreme form of wickedness, according to this [Christian] conception, is exemplified by Satan, who is sometimes conceived of as doing what is morally wrong just because he believes it to be wrong . . . The wickedness of mere human beings seems to result—in normal cases—not from their having a direct desire to do what is morally wrong for its own sake, but rather from their lacking a desire to avoid moral wrongdoing (because, say, they lack a concern for the welfare of others) or from their preferring, say, the pursuit of their own good when these conflict. R. Milo, Immorality (Princeton: Princeton University Press, 1984), p. 7.

Also see W. D. Ross, who says, 'Evil, be thou my good' is the maxim not of a man but of a devil.' W. D. Ross, *The Right and the Good* (Oxford: Oxford University Press, 1930), p. 163.

[21] It is worth explaining why this constraint is not too demanding. After all, one might object that many otherwise moral agents seem to make and be appropriately motivated by their sincere first-person moral judgements even though the content and implication of those judgements remain opaque to them. In fact, on one well-known account of moral judgements, if a person believes that φ-ing in C is right, then she believes that her fully rational self would desire to φ in C, but this could well be the case without her now being consciously aware of that fact.

But the constraint is not motivated by what an agent has to achieve in order to make a genuine moral judgement; it is motivated by what the motivational externalist has to do in order to convince us that the *amoral* agent has made a genuine moral judgement. The challenge must present a cognitively impeccable agent because any linguistic incompetence or a lack of moral understanding on the amoralist's part can be used to explain the lack of appropriate motivation. Most motivational externalists (at least implicitly) recognize this and attempt to abide by this constraint. This certainly explains why most are careful to construct cases in which it is not plausible to claim that the amoral agent is incompetent, morally ignorant or incapable of intelligent discussion about moral matters.

[22] T. Burge, 'Intellectual Norms and Foundations of Mind', *Journal of Philosophy* 83 (1986), pp. 697–720.

[23] T. Burge, 'Intellectual Norms and Foundations of Mind', *Journal of Philosophy* 83 (1986), p. 702.

[24] T. Burge, 'Intellectual Norms and Foundations of Mind', *Journal of Philosophy* 83 (1986), p. 702.

[25] T. Burge, 'Intellectual Norms and Foundations of Mind', *Journal of Philosophy* 83 (1986), p. 707.

[26] Doubting the normative characterization of a concept does not—or not necessarily—undermine one's extensional competence. In order to see this a little more clearly, consider two nearby possible worlds. The first is a world in which beings like us never created furniture, but did create items physically identical to our sofas. These items are treated like religious artifacts. Albert would call these items 'sofas', but so would we. It is just that in this world we would either doubt the normative characterization of the concept in question or notice that the item could also function as an item of furniture. The second is a world very close to ours in which beings like us have created sofa-shaped items for sitting in, but evidently not for religious or aesthetic reasons. Albert would call these items 'sofas', and so would we. It is just that in this world, like our current world, Albert doubts the normative characterization of the concept.

[27] T. Burge, 'Intellectual Norms and Foundations of Mind', *Journal of Philosophy* 83 (1986), p. 703.

[28] Burge notes that '[Albert] is willing to test his hypothesis empirically, and the sociological tests he proposes are reasonable. [Albert] also offers to demonstrate by experiment how the delusive memories are produced. He is sophisticated, and the tests would require elaborate controls. We can even imagine that the theory is developed so as to be compatible with all past experience that might be thought to have falsified his theory. Thus a normal but sophisticated conception of confirmation accompanies [Albert's] unusual theory. We may imagine that if we were to carry out his proposed experiments, [Albert] would come to admit that his theory is mistaken.' T. Burge, 'Intellectual Norms and Foundations of Mind', *Journal of Philosophy* 83 (1986), p. 707.

²⁹ The details that Burge gives suggest that this is the correct interpretation. See endnote 28.

³⁰ T. Burge, 'Intellectual Norms and Foundations of Mind', *Journal of Philosophy* 83 (1986), p. 709.

³¹ Michael Smith makes a closely related argument. M. Smith, *The Moral Problem*, Oxford: Blackwell (1994), pp. 68–71. He also denies that reliably determining the extension of a concept is sufficient for understanding or possessing that concept. He makes this point by way of an argument by analogy. He imagines a person who, despite being blind since birth, has a reliable method of using color concepts. But he tells us that '. . . despite the facility such a blind person has with colour language, many theorists have thought that we should still deny that she possesses colour concepts or mastery of colour terms. For, they say, the ability to have the appropriate visual experiences under suitable conditions is partially constitutive of possession of colour concepts and mastery of colour terms' (p. 69). Smith makes his case by comparing moral concepts to another set of 'queer' concepts—that is, concepts that require a particular response. I take Smith's argument to be a particular instance of my more general point. However, while Smith's argument appeals to what we might think is a special feature of understanding 'queer' concepts like moral concepts or color concepts, my argument simply appeals to a general feature of understanding.

³² R. Shafer-Landau, 'A Defense of Motivational Externalism', *Philosophical Studies* 97 (2000), p. 276.

³³ For example, see C. L. Stevenson, 'The Emotive Meaning of Ethical Terms', *Mind* 46 (1937), pp. 14–31; S. Blackburn, *Spreading the Word* (New York: Oxford University Press, 1984); A. Gibbard. *Wise Choices, Apt Feelings* (Oxford: Clarendon Press, 1990); J. Dreier, 'Internalism and Speaker Relativism', *Ethics* 101 (1990), pp. 6–26.

³⁴ For example, see J. McDowell, 'Virtue and Reason', *Monist* 62 (1979), pp. 331–50; D. Wiggins, 'Moral Cognitivism, Moral Relativism and Motivating Moral Beliefs', *Proceedings of the Aristotelian Society* 91 (1990–1991), pp. 61–85; J. Dancy, *Moral Reasons* (Oxford: Basil Blackwell, 1993); M. Little, 'Virtue as Knowledge: Objections from the Philosophy of Mind', *Nous* 31 (1997), pp. 59–79; D. Bromwich, 'Clearing Conceptual Space for Cognitivist Motivational Internalism' *Philosophical Studies* 148 (2010), pp. 343–67; J. Tresan, 'De Dicto Internalist Cognitivism', *Nous* 40 (2006), pp. 143–65.

³⁵ For example, see T. Nagel, *The Possibility of Altruism* (Princeton, N.J: Princeton University of Press, 1970); J. McDowell, 'Virtue and Reason', *Monist* 62 (1979), pp. 331–50; M. Little, 'Virtue as Knowledge: Objections from the Philosophy of Mind', *Nous* 31 (1997), pp. 59–79; T. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998); D. Bromwich, 'Clearing Conceptual Space for Cognitivist Motivational Internalism', *Philosophical Studies* 148 (2010), pp. 343–67. However, some cognitivist motivational internalists argue that moral judgements are not constituted by motivational attitudes but are necessarily accompanied by them. In particular, see J. Tresan, 'De Dicto Internalist Cognitivism', *Nous* 40 (2006), pp. 143–65.

³⁶ As Mark Van Roojen points out, the desire in question 'must be neither entailed by the presence of, nor be partially constitutive of, any belief.' M. van Roojen, 'Humean Motivational and Humean Rationality', *Philosophical Studies* 79 (1995), p. 37. Michael Smith is the chief contemporary exponent of the Humean theory of motivation. M. Smith, 'The Humean Theory of Motivation', *Mind* 96 (1987), pp. 36–61, and M. Smith, *The Moral Problem*, (Oxford: Blackwell, 1994), especially chapter four.

³⁷ The Humean theory of motivation is controversial. It is rejected by, among others, T. Nagel, *The Possibility of Altruism* (Princeton, NJ: Princeton University of Press,

1970); J. McDowell, 'Virtue and Reason', *Monist* 62 (1979), pp. 331–50; M. Platts *Ways of Meaning* (London: Routledge and Kegan Paul, 1979); D. McNaughton, *Moral Vision* (Oxford: Basil Blackwell, 1988); D. Wiggins, 'Moral Cognitivism, Moral Relativism and Motivating Moral Beliefs', *Proceedings of the Aristotelian Society* 91 (1990–1991), pp. 61–85; J. Dancy, *Moral Reasons* (Oxford: Basil Blackwell, 1993); M. Little, 'Virtue as Knowledge: Objections from the Philosophy of Mind', *Nous* 31 (1997), pp. 59–79; T. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press, 1998); C. Miller, 'Motivation in Agents', *Nous* 42 (2008), pp. 222–66; G. F. Schueler, 'The Humean Theory of Motivation Rejected', *Philosophy and Phenomenological Research* 78 (2009), pp. 103–22; D. Bromwich, 'Clearing Conceptual Space for Cognitivist Motivational Internalism', *Philosophical Studies* 148 (2010), pp. 343–367. Since the Humean theory of motivation gets much of its plausibility from Michael Smith's counterfactual understanding of the direction of fit metaphor (see M. Smith, 'The Humean Theory of Motivation', *Mind* 96 [1987], p. 54 and M. Smith, *The Moral Problem*, Oxford: Blackwell [1994], p. 115), it is also worth noticing that his use of the metaphor has come in for considerable criticism. See, for example, I. L. Humberstone, 'Direction of Fit', *Mind* 101 (1992), pp. 59–83; M. Little, 'Virtue as Knowledge: Objections from the Philosophy of Mind', *Nous* 31 (1997), pp. 59–79; D. Sobel and D. Copp, 'Against Direction of Fit Accounts of Belief and Desire', *Analysis* 61 (2001), pp. 44–53; G. F. Schueler, 'Pro-Attitudes and Direction of Fit' *Mind* 100 (1991), pp. 277–81; S. Tenenbaum, 'Direction of Fit and Moral Cognitivism', in *Oxford Studies in Metaethics*, ed. R. Shafer-Landau (Oxford: Oxford University Press, 2006), pp. 235–64.

[38] See S. Svavarsdóttir, 'Moral Cognitivism and Motivation', *The Philosophical Review* 108 (1999), pp. 161–219.

[39] See S. Svavarsdóttir, 'Moral Cognitivism and Motivation', *The Philosophical Review* 108 (1999), pp. 176–77.

[40] Now, given that we are attempting to figure out Patrick just by paying attention to this behavior, we must rely on the assumption that he has a fairly stable character. Svavarsdottir provides additional information about Patrick's past behavior to support our assumption that he has a stable character:

Additional Information about Patrick: Besides being known for courage and conservative estimates of risk to himself, Patrick is independently minded and earnest to a fault—indeed, honest to the point of tactlessness and even cruelty. And in any case, he has nothing to gain from misleading Virginia in the given circumstances. Moreover, Patrick makes claims couched in moral terms infrequently and impassionately, and never gives them as reasons for his actions. He has frequently been observed taking actions that seem pretty uncontroversially wrong (or in other ways morally problematic) without displaying any signs of regret or shame when his plans have misfired, he has overestimated risk to himself, or he has publicly embarrassed himself in matters he finds important. He has also passed up numerous opportunities to perform obvious and uncostly moral deeds. However, when prodded, he will engage in prolonged and intelligent conversations about moral matters and seemingly take an independent stand on the moral status of a controversial public policy or an action. Nonetheless, he usually ends such conversations by volunteering the opinion that he has long ago rid himself of any aspiration to live by moral standards. S. Svavarsdóttir, 'Moral Cognitivism and Motivation', The Philosophical Review 108 (1999), pp. 177–78.

[41] S. Svavarsdóttir, 'Moral Cognitivism and Motivation', *The Philosophical Review* 108 (1999), p. 178.

⁴² Nick Zangwill makes this point. N. Zangwill, 'Externalist Moral Motivation', *American Philosophical Quarterly* 40 (2003), p. 143. He is also one of the few motivational externalists to provide a positive argument for this position.

⁴³ The thought, or perhaps hope, is that empirical studies on depression and psychopathy might finally prove the possibility of amoralism. See endnote 7.

⁴⁴ M. Smith, *The Moral Problem* (Oxford: Blackwell, 1994), p. 61.

⁴⁵ It is also rather problematic. Mark van Roojen argues that this formulation of the thesis usually involves the assumption that since only desires can motivate action unassisted, all rational persons making sincere first-person moral judgements must also have the desire to act morally (or a desire to act as they judge they must). But, as van Roojen points out, Derek Parfit gives us reason to reject this claim. After all, '. . . some values come only to those who do not desire them. A rational person who believed them valuable would do well to not try to desire them. Thus, a rational person would not always desire to have or do what she believes valuable' (p. 27). See M. van Roojen, 'Humean and Anti-Humean Internalism about Moral Judgements', *Philosophy and Phenomenological Research* 60 (2002), pp. 26–49. Also see D. Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), especially Part I.

⁴⁶ Some motivational externalists formulate a stronger version of the thesis designed to be incompatible with a particular version of motivational internalism. For example, Nick Zangwill states that 'Motivational externalism is the view that moral judgements have no motivational efficacy in themselves, and that when they motivate us, the source of motivation lies outside the moral judgement in a separate desire.' N. Zangwill, 'Externalist Moral Motivation', American Philosophical Quarterly 40 (2003), p. 143. This thesis is incompatible with the view that moral judgements are at least partially constituted by motivational attitudes. However, and as noted in endnote 5, motivational internalism is just the thesis that there is a necessary connection between moral judgement and motivation. It is therefore compatible with the view that moral judgements are not constituted by motivational attitudes but are necessarily accompanied by them. For examples, see J. Tresan, 'De Dicto Internalist Cognitivism', Nous 40 (2006), pp. 143–65; C. Strandberg, 'The Pragmatics of Moral Motivation', Journal of Ethics 15 (2011), pp. 341–69; and F. Bjorklund, G. Bjornsson, J. Eriksson, R. Francen Olinder & C. Strandberg, 'Recent Work on Motivational Internalism', Analysis 72 (2012), pp. 124–37.

That said, many motivational internalists hold that sincere, first-person moral judgements necessarily motivate those who make them *because* moral judgements are at least partially constituted by motivational attitudes. For these internalists, the move toward a defeasible conceptual connection between moral judgement and motivation is especially problematic. After all, once the *ceteris paribus* clause is inserted between moral judgement and motivation, it is clear that factors external to the moral judgement are necessary for motivation.