

## Chapter 37

### Motivational Internalism and Externalism

G. F. SCHUELER

The internalist-externalist debate in moral philosophy is about the place of motivation in morality and in moral judgments (where ‘motivation’ can refer either to having a normative reason to do something or actually being inclined to do it). The underlying issue is about the connection between motivation and judgment generally. Intuitively, moral requirements or prohibitions give those subject to them reason to act in the way required or to refrain from acting in the way forbidden. Analogously, the belief that one is required to do something, or prohibited from doing it, should move one to act or refrain from acting in the designated way. These are internalist intuitions.

Broadly speaking, internalists hold that moral judgments (or alternatively moral ‘facts’ such as obligations) always somehow involve the agent having some motivation, either toward the thing positively evaluated or away from the thing negatively evaluated. Externalists deny these things. Part of the interest in these opposed positions is that, while apparently incompatible, each seems to be supported by a very plausible meta-ethical intuition. On one side, moral judgments, especially ones about one’s own actions, seem paradigmatically practical. Someone who blithely and without hesitation did something she claimed to think morally wrong would be regarded as insincere in her claim, not really thinking it wrong at all. On the other side, the most straightforward form of moral realism, the sort that holds that moral wrongness, say, is simply a property of some actions, seems to entail that one could realize that an action one was considering had that property while being unmoved by that fact, just as one might be unmoved by any other fact about it.

This debate over the role of motivation in moral judgments is strongly influenced by the wide acceptance of the Humean Theory of Motivation. According to this view, there is a sharp difference between beliefs and desires (or states with ‘mind-to-world direction of fit’ and those with ‘world-to-mind direction of fit’). And on this theory, all motivation requires desire. Beliefs alone cannot motivate. So if one accepts this theory one seems committed to one of two positions, each of which seems problematic. Holding the ‘cognitivist’ view that moral beliefs have truth values like any other beliefs seems to commit one to the externalist position that moral beliefs involve no motivation by themselves. That makes moral motivation dependant on an extra, apparently non-rational, desire to do what is right (or the like). Alternatively, if one wants to be an internalist about moral judgments, one seems committed to holding that what are called ‘moral beliefs’ are really some form of desire or ‘pro-attitude’, which makes one a ‘non-cognitivist’ about moral beliefs and seems to entail that such ‘beliefs’ cannot have truth values. So trying to combine these two initial intuitions, that moral beliefs both motivate and have truth values like any other beliefs (i.e. ‘cognitivist internalism’) seems to require rejecting the Humean Theory of Motivation.

There is yet another complicating factor. It might be thought that motivation takes place via reasons, or the agent’s conception of her reasons. And one could be an internalist or externalist about reasons, or reason judgments, as well as about moral facts or judgments. So there are two further issues. First, if one thinks that motivation takes place via reasons for acting (or the agent’s judgment about this) that requires understanding the question of whether moral judgments (or facts) motivate in terms of the question of whether such judgments or facts provide reasons to act. Second, since reasons, or reason judgments, also seem to move us, the same issue about the connection with inclination to act arises for reason judgments (or facts) as arises about moral judgments (or facts).

To distinguish between whether we are speaking of judgments or the ‘facts’ these judgments are about, we can follow Darwall and distinguish ‘judgment internalist’ from ‘existence internalist’ views. Judgment internalists “hold it to be a necessary condition of a genuine instance of a certain sort of *judgment* that the person making the judgment be disposed to act in a way appropriate to it.” Existence internalism, however, “is the thesis that something is a ground of an act’s actually having a certain property (being right, for example ...) only if it is capable of motivating the agent.” (Darwall, 1983, p. 54) If we focus first on moral claims, the distinctions between moral judgments and moral facts, and between reasons and motives, mark out four possible versions of internalism. (This part of the taxonomy of varieties of internalism to be explained below follows Russ Shafer-Landau.) (Shafer-Landau, 2003, Chapter Six.)

There is, first, the sort of existence internalism that claims a necessary connection between moral facts (such as obligations) and justifying reasons. That would seem true for instance if ethical egoism is the correct moral theory, since egoism says that what one morally should do is what will serve one’s own interest and, presumably, one always has good reason to do what is one’s own interest.

Second, there would be the sort of existence internalism that held that the mere fact of the existence of a moral requirement is necessarily connected to the motivation of the agent under the requirement. This looks plausible if one holds that moral requirements are not independent of the attitudes of the agent, e.g. that it is a necessary condition of someone being morally required to do something that she have a certain attitude toward that thing. This later becomes plausible if e.g. one combines the view that moral facts entail justifying reasons (the first view just distinguished) with the idea that justifying reasons themselves entail some sort of internal motivating state in the

agent (which is an existence internalist view about reasons, not moral facts). Bernard Williams famously argues for this latter view. (Williams, 1981)

The other two sorts of ‘ethical’ internalism start with moral judgments, rather than moral requirements or other moral facts. So a third variety of internalism claims that there is a necessary connection between moral judgments and justifying reasons, more specifically that having a good reason for doing or avoiding something is a necessary condition of judging some act to be morally required or prohibited. This form of internalism would be entailed if one held both that moral judgments entail motivation to follow them (i.e. the fourth variety below) and that it is a sufficient condition of having a reason to act that one have some motive to do so.

The fourth variety of internalism holds that there is a necessary connection between moral or evaluative judgments and the motivation of the agent making the judgment. Emotivism, for instance, is this sort of judgment internalism (as are I think its contemporary, ‘expressivist’ descendants).<sup>1</sup> It claims that to make a moral judgment just is to express a certain attitude. And since emotivists hold that attitudes are motivating states, on this view making a moral judgment involves, among other things, being inclined or motivated toward the thing being evaluated, if the judgment is positive, or against it if the judgment is negative.

If this kind of view were correct it would solve (or ‘dissolve’) the apparent puzzle about the connection between moral facts and agent motivation in a very direct way. The apparent motivational ‘pull’ (or ‘push’) of things with evaluative properties would be an artifact of this feature of moral or evaluative judgments, rather than being a feature of the things themselves. According to this sort of view we only judge things to have such properties if we have some motivating attitude toward them.

As it stands though this taxonomy of internalisms is incomplete since it leaves out internalism about either reasons or reason judgments. Bernard Williams' 'internal reasons' view is of the former sort. According to Williams (roughly) one only has a good reason to do something if that thing promotes the satisfaction of one of one's pre-existing desires (or other 'subjective motivational' states). (Williams, 1981) This is a form of existence internalism about normative reasons rather than about any specifically moral fact such as a moral obligation. If one accepts the first variety of internalism above then the notion of an agent having a good reason to do something can provide a 'middle term' connecting claims about moral requirements and the actions that constitute following those requirements. Michael Smith, for instance, argues that a form of judgment internalism that claims a connection between moral judgments and motivation (the fourth variety above) is entailed by a form of existence internalism that connects moral facts with reasons (the first variety above). This is because, he says, "an agent who judges herself to have a reason to act in a certain way ... is practically irrational if she is not motivated to act accordingly. For if she is not motivated accordingly then she fails to be rational by her own lights." So Smith holds that the first variety entails the fourth variety because, "according to the [first variety], the judgement that it is right to act in a certain way is simply equivalent to the judgement that there is a reason to act in that way." (Smith, 1994, p. 62)

Williams' view also makes sense of (though it does not fit into) the second of the four categories just distinguished, the one that would make it a necessary condition of having a moral obligation that the agent have some motive to act on that obligation. Again, it does this if we accept the view delineated in the first category, that it is a necessary condition of having a moral obligation that one have some good reason to act on it. If this is true then if Williams is right and it is a necessary condition of having such a reason that one have some pre-existing motive, it follows that it is a necessary condition of having a moral obligation that one have some pre-existing motive which would lead one to act on it.

But this fifth form of internalism, exemplified by Williams' internal reasons doctrine, though it connects reasons and motives, leads to a form of skepticism about morality. If one accepts it then one is faced with an unhappy choice that turns on one's view about the first kind of internalism. On the one hand, one can deny this kind of internalism, i.e. deny that obligations and other such 'moral facts' entail reasons to act on them. That would mean however that even though Gyges, in Plato's famous story, may have been under a genuine moral requirement not to do the immoral things his invisibility allowed him to do, still he may have had no actual reason to act on this requirement, if there was nothing that moved him to do so. Alternatively, if one wants to hold that moral requirements do entail reasons, then accepting Williams' internal reasons claim means that Gyges may have been under no actual moral requirement to forgo what he did, i.e. if he had no inclination to do so.

The plausibility of the fourth sort of internalism, the judgment internalist idea that there is a necessary connection between an agent's judgment that she has a moral obligation (say) and being moved to act on that judgment, rests at least partly on the intuition mentioned above that someone who makes such a moral judgment but has no inclination to act on it would be revealed as insincere. A similar intuition holds for reason judgments and motivation. Someone who agrees that she has good reason to do something while remaining utterly unmoved to do it even when she thinks she has no reason not to, looks insincere in her agreement. And as Smith's argument suggests, it is plausible to think that the first intuition rests on the second, since the connection between moral facts and reasons (the first sort of internalism) is itself widely accepted. So it may help to evaluate at least some of the versions of ethical internalism if we look more carefully at the reasons internalism that connects reason judgments and motivation (which is the sixth variety to be distinguished). In the terminology adopted above, this is judgment internalism about the agent's own normative or justifying reasons.

The claim to examine is that

[JIR] *it is a necessary condition of making a genuine judgment to the effect that one has a good reason to do something that one be at least to some minimal degree motivated or inclined to do it.*

The plausibility of this claim rests on the fact that it seems to provide something close to the minimum connection between the normative and the practical. Even if one doesn't think that obligations or reasons connect to motivation, or that ethical judgments automatically motivate, it seems hard to deny that judgments about *their own reasons* motivate those who make them, at least if they are rational. That seems to be what Smith has in mind when he writes, in the argument quoted above, that "an agent who judges herself to have a reason to act in a certain way ... is practically irrational if she is not motivated to act accordingly." If [JIR] is not true then it seems very difficult to explain what the connection is between the agent's own view of the reasons she has for doing something and her actually doing it. And it seems that there must be some connection. Otherwise judgments about one's reasons to act will make it no more likely that one actually will perform the action than will one's judgments about, e.g., whether the action will happen on an even-numbered day of the month.

What is it to make a judgment that one has good reason to do something? Expressivist and similar speech-act accounts of judgments, to be plausible at all, will have to make judgments into silent, inner speech, since not all judgments are in the form of actual utterances. But of course I can say things, even only to myself, which I know are false. And similarly most of the things I judge (believe) to be true I am not even thinking about at the moment, let alone talking about (even to myself). These are familiar points. But they apply to the judgments one makes about one's own reasons as much as to any other judgments. Sometimes, when I am trying to

figure out what to do, I explicitly consider the various things that I take to count for or against the action I am considering. But such explicit deliberation is rare, and for the same reason explicit rehearsal of ones beliefs is rare. Life is just too short and our beliefs far too numerous. Plus, since deliberation is itself something one can have reasons for and against doing, it is itself something one could deliberate about doing. So, on pain of regress, it is not possible that all the things one does for reasons are things one deliberates about.

There are really two distinct questions: What is it to act for a reason? What is it to judge that one has a reason to do something? [JIR] is about the second, but the first is the more general question. We can make a start on seeing what the connection between the two is if we make the simplifying assumption that the agent in some case does explicitly deliberate and then acts on the basis of her deliberation. Consider such a case.

A few moments ago I got up from my desk and walked to the kitchen to get some coffee. Suppose that when I reach the point in the hallway where I need to turn left in order to get to the kitchen, I pause and consciously try to figure out what to do.

‘Lets see,’ I say to myself, ‘I want to get some coffee and at this point in the hallway there are four possible ways to go.

(1) I could continue down the hallway in the same direction I am now heading. That will take me into my wife’s workroom.

(2) I could turn left, which is the direction of the kitchen, where the coffee is.

(3) I could turn around and go back to my desk.

Or (4) I could just stand here in the hallway.

Going to my wife’s workroom is not a bad idea, since it is always nice to see what she is working on, but that might distract me from what I am doing and anyway I want to get some coffee at the moment and there is none in her workroom. There is nothing to be



said for just standing in the hallway or turning back toward my desk in defeat. The coffee, which is what I am after, is in the kitchen to the left.

So, I'll turn left.'

Notice that, though these judgments set out my reason for turning left, nowhere in this deliberation does the claim that I have reason to turn left appear. So did I, just as I turned, judge that I had good reason to turn left? I don't think it is obvious that the answer to this question is 'yes', even though I clearly had a reason for turning at that point and did indeed turn for that reason.

Someone might think that the above description just leaves out the relevant reasons claim. And of course I might have made that claim while deliberating. So the question is really whether without such a claim my deliberation would have been incomplete or defective. I think the answer is that it would not be. What gives me reason to turn left is the fact that what I want, the coffee, is to the left. So how could there be anything wrong with the reasoning in which I take account of this fact and, as a result, decide to turn left?

Lets assume that, given that I want to get some coffee and the coffee is in the kitchen to my left, I have reason to turn left. That truth is not part of my reasoning because it cannot be one of my reasons. What has to be part of my reasoning is what makes it true that I have reason to turn left, i. e. the fact that what I want, the coffee, is to the left. Here is a different sort of example. Suppose I believe that P and that if P then Q. Assuming that these two beliefs are themselves well founded, I then have good reason to believe that Q. Does that mean that, really, in order to be completely accurate, all reasoning via modus ponens should be represented as follows?

- (1) P
- (2) If P then Q
- (3) Premises (1) and (2) provide good reason to believe Q
- (4) So, Q

No. As Lewis Carroll showed a long time ago, this immediately leads to a regress. (Carroll, 1995, 691-693) Exactly the same bad argument that seems to require the addition of (3) leads to the addition of another premise to the effect that (1) through (3) provide reason to believe Q. And so on. The sentence labeled (3) here is not part of the reasoning at all. That does not mean that someone might not utter (3) in the course of her attempt to figure out whether Q is true. It means that the argument here is a perfectly good one (logically valid in fact) without the addition of (3).

Exactly parallel considerations apply to practical reasoning, and so to whether we need to add the claim that I have reason to turn left to the description of my deliberation in the hallway. If the fact that I have a reason is itself a reason, on a par with the fact that the coffee I want is in the kitchen to the left, then my deliberation won't be correct until I include this further reason as one of my reasons. And then we get the same issue for that addition, and a similar regress.

The claim that some fact gives someone a reason to do something is a claim about the rational status of that action with respect to that fact. It is not, therefore, part of what makes it reasonable to perform that action. So the description above of my deliberation is acceptable as it stands. Though I might have had the thought that I have good reason to turn left, that is not part of my deliberation. It is a comment *about* my reasons, not one of the reasons themselves.

[JIR] is the claim that it is a necessary condition of an agent making a judgment that she has reason to do something that she be at least to some minimal extent motivated to act on that reason. That means that a good place to look for such judgments is when agents are motivated to act, and in particular when they do act for reasons. But when we examine the example of my left turn toward the coffee in my kitchen it turns out that I need make no actual reason judgment. In fact, the thought that I must employ such a judgment generates a regress. Yet I still have a reason to act and in this example I do indeed act for that reason.

This does not show that [JIR] is false of course. I have been arguing that one can act for a reason without making the judgment that one has a reason, not that one can make such a judgment without being moved to act. But this argument does suggest that even if this form of internalism is true it doesn't really help us make sense of what it is to act for reasons. It also raises the question of what the supposed 'motivation' claimed by this form of judgment internalism could possibly be. If the argument above is correct then it could not be the sort of motivational connection represented in practical deliberation when someone considers what reasons she has and then acts on this deliberation. Judgments that I have a reason are not part of practical deliberation at all. They are descriptions of the status of one's reasons. So if such judgments somehow motivate actions, this must be some sort of extra- or non-rational form of motivation.

[JIR] is important because it links moral judgment and motivation, and it seems plausible because it seems to show that we can make sense of why people are moved to act on the reasons they judge they have by the idea that it is a condition of even making such a judgment that one have some inclination to act in the way specified. But if [JIR] is to help us understand what it is to act for a reason, we must presuppose that when someone acts for a reason she judges that she has a reason to act in the way she does. As we have seen, this is not the case.

*See also* PRACTICAL REASONING (18), MOTIVATING REASONS (25), HUMEAN THEORY OF MOTIVATION (26), RATIONALITY (45), SPEECH ACTS (10)

### References

Simon Blackburn, *Essays in Quasi-Realism* (Oxford University Press, 1993).

Bernard Williams, "Internal and External Reasons", reprinted in his *Moral Luck* (Cambridge University Press, 1981)

Lewis Carroll, "What the Tortoise Said to Achilles", *Mind*, Vol. 104, No. 416 (October 1995), pp. 691-693, reprinted from *Mind*, 4 (1895).

Stephen Darwall, *Impartial Reason*, (Cornell University Press, 1983)

Russ Shafer-Landau, *Moral Realism*, (Oxford University Press, 2003), Chapter Six

Michael Smith, *The Moral Problem* (Basil Blackwell, 1994)

Charles Stevenson, *Facts and Values* (Yale University Press, 1963)

---

<sup>1</sup> The classic expression of emotivism can be found in (Stevenson, 1963). For an example of contemporary expressivism see virtually any of the essays in (Blackburn, 1993).