The argument for internalism: reply to Miller

MICHAEL SMITH

I hold that the following is a conceptual truth (Smith 1989; Smith 1994: 60–76):

Weak Internalism: If an agent judges it right to φ in certain circumstances C, then she is motivated to φ in C, at least absent weakness of will and the like.

This is a weak principle because those who accept it can agree that agents who judge it right to act in a certain way may not be motivated to act in that way at all if they are weak-willed. And even when such agents are not weak-willed, and so are motivated accordingly, those who accept this principle can agree that their motivation may be overridden by other motivations such agents may have.

In The Moral Problem (1994) I provide an argument for this weak form of internalism and against the corresponding form of externalism. The argument begins from a premiss acceptable to both internalists and externalists alike. The premiss states that it is a striking fact that when good and strong-willed people judge it right to φ in C they are motivated to φ in C. Changes in the moral judgements that good people make bring corresponding changes in their motivations, at least absent weakness of will and the like, and necessarily so. The argument I give for internalism is that it can, whereas externalism cannot, provide a plausible explanation of this striking fact.

Alexander Miller objects to my argument (ANALYSIS, this issue). He points out that either the striking fact that internalism can explain and externalism cannot is a conceptual truth, or it is merely an empirical truth. But if it is a conceptual truth then it simply amounts to a statement of internalism itself, and then the fact that externalism cannot explain it hardly counts as an argument in favour internalism over externalism. The argument begs the question against externalism. And if the striking fact is merely an empirical truth then it is false that internalism can explain it. Internalism is, after all, supposed to be a conceptual truth, and so is inconsistent with the claim that the striking fact expresses merely an empirical truth, given that the striking fact is simply a statement of internalism itself.

Miller is right that the striking fact must express a conceptual truth for my argument to succeed. That is how I intended it to be taken. But he is wrong that if it expresses a conceptual truth then my argument somehow

References


3 How might Smith attempt to respond to this argument? One way would be to advert to the responses that have been developed to Mark Johnston’s ‘missing-explanation’ argument (cf. Johnston 1993) in the recent literature (cf. Menzies and Pettit 1993). Recall that Johnston’s argument is that the claim that our concept of red is the concept of being disposed to look red to normal observers in standard conditions rules out the possibility of there being an empirical explanation of the following form: an object looks red to a normal observer in standard conditions does so because it is red. Now I have myself developed some responses to Johnston’s argument (1995, 1996 forthcoming) but I cannot see how those responses, or the other responses in the literature, can be invoked to defend Smith against the argument in the text. Firstly, the conceptual claim at the centre of dispositionalist theories of colour, namely that our concept of redness is the concept of a disposition to appear red to normal observers in standard conditions, has a different form from the central internalist claim, that it is a conceptual fact that if an agent judges that it is right to G in circumstances C, then absent practical irrationality, she will be motivated to G in circumstances C. Secondly, and more importantly, the responses to the missing-explanation argument attempt to establish only that the conceptual claim at the heart of dispositionalism about colour does not rule out an empirical explanation of the relevant sort: they do not purport to establish that such an empirical explanation could be provided merely by citing the dispositionalist’s conceptual claim. So I cannot see how the available responses to the missing-explanation argument, my own included, could possibly be of any help to Smith here.

4 Thanks to Nick Dent, John Divers, Brian Garrett, Duncan McFarland, Stephen Mumford, Laura Schroeter, Michael Smith, Jim Stuart, Mark Walker and Jo Wolff for comments and discussion.
begs the question against externalism. On the contrary, the beauty of the striking fact is precisely that internalists and externalists alike must accept its status as a conceptual truth. However, the very fact that a clear-headed philosopher like Alexander Miller should think otherwise makes me worry that my argument is being misunderstood more broadly. I would therefore like to take the opportunity here to set the argument out once again, and to make it perfectly plain why it does not beg the question.

As I said, I believe Weak Internalism expresses a conceptual truth. Externalists claim it does not because, as they see things, it is impossible for there to be people who judge it right to φ in C without being motivated to φ in C and without suffering from weakness of will. Indeed, externalists like David Brink claim that we have a name for such people (Brink 1986): we call them amoralists. Let's grant this definition of 'amoralism', for the sake of argument.¹ That is, let's assume, with the externalists, that the following principle:

Amoralist Externalism: If an agent judges it right to φ in C, and she is an amoralist, then, even if she does not suffer from weakness of will and the like, she may not be motivated to φ in C.

provides us with a definition of the term 'amoralist': an amoralist is simply someone who makes Amoralist Externalism come out true. If it is so much as possible for there to be any amoralists, then we are defined by Amoralist Externalism, then Weak Internalism is false. The argument for Weak Internalism I offer in The Moral Problem is, in essence, that it is impossible for there to be any amoralists. Here is why.

Note that just as we can define what an 'amoralist' is in terms of Amoralist Externalism, we can similarly define the contrast class of 'moralists': moralists are those people who are such that, when they make judgements about what it is right to do they are motivated to act accordingly, at least absent weakness of will and the like. The moralists' possession of this disposition is what marks them off from the amoralists. Externalists are therefore committed to the following principle:

Weak Moralist Internalism: If an agent judges it right to φ in C, and that agent is a moralist, then she is motivated to φ in C, at least absent weakness of will and the like.

They are committed to this principle's being a conceptual truth by their definition of the 'amoralist' and the fact that we use that definition to define the contrast class of moralists.

Even though moralists may have true or false moral beliefs it seems plain that they possess an executive virtue that amoralists simply do not possess:

¹ Though see footnote 2 below.
marks nothing in reality: it marks nothing in reality because everyone capable of making a moral judgement is a moralist. Moreover – and, as we will see in a moment, importantly for present purposes – internalists are therefore bound to insist that the connection between judgements of fundamental rightness and motivation has a very distinctive characteristic.

Suppose someone judges it right to φ in C, but not in virtue of any further right-making feature of φ-ing. They judge it fundamentally right to φ in C, simply as such. For example, suppose that the person is a utilitarian and that he judges it right to maximize happiness and minimize suffering. The internalist is bound to insist that the motivation this agent must have, if he makes this judgement and is not weak-willed, is non-instrumental in character: that is, the utilitarian’s motivation to maximize happiness and minimize suffering is not dependent on the fact that its satisfaction contributes towards the satisfaction of some further motivation that the agent has. The maximization of happiness and minimization of suffering is the ultimate consideration that moves the agent. This is a direct corollary of the fact that the judgement concerns fundamental rightness, according to internalists.3

Now consider the externalists’ explanation of the truth of Weak Moralist Internalism. Externalists think that the distinction between moralists and moralists does mark something in reality. They therefore think that both facts make a contribution to the explanation of the consequent: it is in part because of the nature of the judgement that the moralist makes – the fact that the judgement is a judgement about rightness – and it is in part because of the nature of the moralist herself – the fact that she has the nature she has as opposed to the nature of an amoralist – that a moralist is either motivated or weak. But why? What is the difference supposed to be between the nature of the moralist and the nature of the amoralist? And how do these facts combine to explain the consequent?

---

2 However, since the term ‘amoralist’ is a term that has an ordinary usage, and since we ordinarily take there to be exemplars of amoralism as that term is ordinarily used, it follows that if the externalists’ definition of ‘amoralism’ makes the class of amoralists necessarily empty then we should look elsewhere for a definition of ‘amoralism’ that gives the meaning of the term as it is ordinarily used. For a discussion of alternative internalist definitions of ‘amoralism’ see Smith 1994: 67–68.

3 In support of this claim, consider the two standard ways in which internalists account for the internalist nature of moral judgement. Internalists of an anti-rationalist persuasion – those who are expressivists, say, and so think that such judgements are expressions of desires or dispositions to have desires – claim that the judgement that it is right to maximize happiness and minimize suffering, concerned as it is with fundamental rightness, is an expression of a non-instrumental desire to maximize happiness and minimize suffering, or perhaps an expression of the agent’s disposition to have such a non-instrumental desire when not weak-willed. The desire is non-instrumental in character because its satisfaction will not serve any further desire that the agent has. Internalists of a more rationalist persuasion – those who adopt a position more like the one I argue for in The Moral Problem – claim that the judgement is an expression of the agent’s belief that she would have a non-instrumental desire that she maximizes happiness and minimizes suffering if she were fully rational, and that, because she has this belief, she must either desire non-instrumentally to maximize happiness and minimize suffering or else suffer from a certain sort of incoherence, which is what weakness of will turns out to be, on this sort of view.
The only answer available to the externalist would seem to be that what makes someone a moralist, as opposed to an amoralist, is the fact that she contingently desires to do what is right. Amoralists, by contrast, contingently desire no such thing. This is why both facts make a contribution to the explanation of the consequence of Weak Moralist Internalism. It is in part because a moralist desires to do what is right, and in part because she believes that it is right to φ in C, that she either desires to φ in C or else suffers from some form of weakness; this is why changes in her moral motivations follow reliably in the wake of changes in her judgements. Her desire to φ in C is derived instrumentally from her non-instrumental desire to do what is right together with her means-end belief that she can do what is right by φ-ing in C. The only thing that would block this transition is a failure of instrumental rationality. That is what weakness of will amounts to, on this sort of view.

But let’s now think more carefully about the explanation of Weak Moralist Internalism we have been offered. The externalist’s idea is that anyone who is a moralist — that is, anyone whose motivations track their moral judgements in a reliable way — must have as their primary source of motivation — the motivation that drives all of the rest — a desire to do what is right. I have two main objections to this. The first is that if we accept the externalists’ idea then we must redescribe familiar psychological processes in ways that depart radically from the descriptions that we would ordinarily give of them. The second is that if we accept the externalists’ idea then we thereby commit ourselves to an implausible conception of moral perfection. Externalism is therefore a radically revisionary theory. Let me take these two points in turn.

Here is a story about a friend of mine. When he was a beginning graduate student he came under the influence of Peter Singer’s writings and converted to utilitarianism. He believed that it is always right to maximize happiness and minimize suffering, and fundamentally so: he didn’t believe that acts with this feature were made right in virtue of their possession of some further right-making feature. Over the years, however, via a process of argument, he came to change his mind. He now believes that it is sometimes right to give extra benefits to his family and friends, even when doing so cannot be given a utilitarian justification. He is also a moralist. His moral motivations have followed reliably in the wake of his new judgement, at least absent weakness of will.

But when I reflect on my friend’s change of mind, and the change of heart his change of mind has caused in him, it seems to me that he has changed from being a utilitarian monster, someone whose only non-instrumental desire was to maximize happiness and minimize suffering, and that what he has changed into, as a result of changing his moral beliefs, is someone who also has more everyday non-instrumental desires, desires to confer special benefits on his family and friends and the like. In other words, it seems to me that what he has acquired are new non-instrumental personal concerns, whereas before he had only one concern: a non-instrumental impersonal concern.

Externalists insist that this description of my friend’s psychology — the description that he is disposed to give, that I am disposed to give, and that all his other friends and family are disposed to give as well — is thoroughly mistaken, however. Since he is a moralist externalists tell us that he must really only have had an instrumental desire to maximize happiness and minimize suffering in the first place, and that he must really only have acquired instrumental desires to confer special benefits on his family and friends, at least in the first instance. These desires can all only have been instrumental because, possessed as they are by a moralist, they must all have been derived from the one and only non-instrumental desire he has in virtue of being a moralist: a non-instrumental desire to do what is right that leads him to have various instrumental desires — before, an instrumental desire to maximize happiness and minimize suffering, and now, in addition, instrumental desires to confer special benefits on family and friends as well — when combined with his earlier and later means-end beliefs about how to do what is right. But this seems to me to be manifestly false, an utterly theory-driven redescriptions of my friend’s psychological change.

If I am right that the redescriptions of my friend’s psychology that the externalists recommend is entirely driven by their commitment to externalism over internalism then, since internalism is consistent with the description that people who are familiar with him are ordinarily disposed to give of his psychological change, it seems to me that we should therefore make the opposite inference instead. Since my friend clearly has been transformed from being someone with one set of non-instrumental concerns into someone with another set, and since the only theory about the nature of a moralist that is consistent with this familiar description of his psychological change is internalism, it follows that we should reject externalism in favour of internalism.

My second main objection to the externalists’ claim that moralists are, by definition, people who have as their single primary source of moral motivation a non-instrumental desire to do the right thing is that it commits them to a quite incredible picture of moral perfection. For ask yourself where someone who is morally perfect fits into the externalists’ picture of the divisions among agents. Are they to be found among the amoralists? Or are they to be found among the moralists? The obvious answer for the externalists to give is that morally perfect agents are to be found among the class of moralists: indeed, they are to be found among the
perfect people, in a quite straightforward sense, alienated from the features of their acts that make them right. In desiring to do what is right for the sake of its being the right thing to do, rather than for the sake of the feature that makes it the right thing to do, they desire something that is not of primary moral importance. They seem precious, overly concerned with the moral standing of their acts when they should instead be concerned with the features in virtue of which their acts have the moral standing that they have. In a word, they do not seem to be morally perfect at all. As I say in The Moral Problem, externalists thus seem to give a moral fetish the status of the one and only moral virtue.

Externalists therefore force us to revise our understanding of moral perfection in intolerable ways as well. And since the revision is once again driven by the externalists’ commitment to externalism over internalism, and since internalism leaves our ordinary conception of moral perfection intact, it seems to me that we should therefore make the opposite inference instead. Since people who are morally perfect clearly are motivated by the very features of their acts that make their acts right, and since internalism is the only theory about the nature of moral judgement consistent with this fact, it follows that we should reject externalism in favour of internalism.

To sum up: Weak Moralist Internalism expresses a conceptual truth that must be accepted by internalists and externalists alike. Since both must accept it, it follows that they must each tell us their very different reasons for accepting it. But when we hear what they have to say it becomes plain that the externalists’ explanation is implausible. Externalists have to say that what makes someone a moralist is the fact that they possess a non-instrumental desire to do the right thing. But this leads them to tell revisionary stories about the nature of familiar psychological processes and the nature of moral perfection, stories that externalists themselves should wish to avoid having to tell. The internalists’ explanation of Weak Moralist Internalism by contrast – the fact that it is entailed by Weak Internalism – has no such revisionary consequences. This is the argument for internalism I gave in The Moral Problem.5

4 I have often been asked whether utilitarians would accept this. The problem is supposed to be that, according to utilitarians, people should have those motivations which are such that, by having them, they lead a life that is utility maximizing. It is therefore a matter to be determined by any particular person’s contingent circumstances whether that person should be motivated to maximize utility. This is supposed to show that utilitarians are committed to rejecting the ordinary conception of moral perfection outlined in the text. But I disagree. It shows at most that utilitarians who accept this ordinary conception of moral perfection are thereby committed to the view that a morally perfect utilitarian, in this ordinary sense, might none the less be morally required to cause himself to be less than morally perfect; that only by becoming less than morally perfect in the ordinary sense can he have the motivations his theory tells him he ought to have. In this connection it is worthwhile re-reading Adams’s classic paper ‘Motive Utilitarianism’ (1976), for he there struggles with just this point during his discussion of conscience-utilitarianism.

5 I am very grateful to Jerome Schneewind, and to the members of his Spring 1996 graduate seminar on ethics at Johns Hopkins University, for their very helpful comments on an earlier presentation of these ideas.
Voting and motivation

Keith Graham

What are people doing when they vote, and how much does it matter? Some may vote for their own interests and some for the common good. Jonathan Wolff (1994) has recently argued that the defence of democracy is greatly complicated by the fact that this mixture of motives may be present. In this he echoes closely an argument originally put forward and recently reiterated by Brian Barry (1990). On the other hand, Andrew Brennan and Loren Lomasky (1993) have suggested that there are grounds for regarding voters as expressing their moral beliefs rather than their own interests, but with similarly troubling results for democracy.

My aim here is not to question whether there is a clear and sound distinction between interests and moral values (though we might). It is rather to suggest that the question whether voters express the first or the second is of less importance than often thought. Democracy may be in need of defence, but that fact has no special connection with the expressive status of votes.

1. Wolff argues that where the mixture of motives is present it is possible to arrive at a majority decision 'which is neither in the majority interest, nor believed by the majority to be for the common good' (Wolff 1994, 194). For example, suppose that A is in the interests of 40% and B in the interests of 60%; that 80% believe B to be the common good (defined as what is in the majority's interests) and 20% believe this of A; and that while those with an interest in A vote according to interest, those with an interest in B vote according to the common good. Then 52% will vote for A, even though it is in the minority interest and believed by only 20% to be the common good. Wolff concludes that 'we cannot rely on any assurances that democratic decision-making reveals either the majority interest or the common good', and that defenders of democracy 'for the time being at least, would do well to downplay its supposed instrumental advantages' (Wolff 1994: 195).

Barry invites us to imagine a situation where some good is to be distributed to A's and B's, and the question is whether the A's should get more of it or the same amount as the B's. Imagine further that the A's outnumber the B's, the A's think that equal distribution would be fair, and the B's are equally divided on that question. Barry concludes:

If everyone voted according to private wants, the A's would win; if everyone voted according to his attitude based on a public standpoint, the B's would win; but if some vote on one ground and some on another the result is completely indeterminate and depends on how many there are of each and how the numbers are composed. (Barry 1990: 64)

For Barry, there is a connection with acceptance of the majoritarian principle. Its most objectionable feature is that 'it commits one to handing over questions of right and wrong, justice and injustice, to the majority of a group in which one's own voice counts only as one'. This is avoided if we take account only of 'desires which are put forward simply by people as wants in matters affecting themselves' (Barry 1990: 63). These carry 'a certain automatic claim to satisfaction' whereas judgements about what is right or fair are simply of a different kind, and compensating people for the failure to satisfy them is inappropriate. Accordingly,

... it is far easier to swallow the 'democratic principle' when the issue at stake can be presented as a conflict of wants than when it appears as a conflict of opinions as to what is right, fair, etc. (Barry 1990: 63 n. 3)

I do not deny the mathematically-based point being made in either case, but I do want to challenge the idea that the untoward results arise from the mix of motivations leading people to vote in each example. In Wolff's case the result obtains because of the possibility of error in judgement. As he himself recognizes, 'the key assumption is that morally motivated individuals can make a mistake about what morality requires' (Wolff 1994: 194).

---

1 This is a revised edition of Barry (1963), with a new introduction. Arabic numeral page references are identical in the two editions.

2 Presumably for Barry 'winning' means receiving the greater share of the good in question. It is already guaranteed that the A's will win in the sense of getting their way since they are stipulated to be in the majority.