Michael Smith

Daniel Star
Boston University

Michael Andrew Smith was born in Melbourne on July 23, 1954. He studied for both a BA and an MA at Monash University, finishing there in 1980. His subsequent studies at the University of Oxford led to him being awarded a BPhil in 1983 and a DPhil in 1989. As a graduate student in Oxford, he was supervised by Simon Blackburn. Smith taught philosophy at Wadham College in Oxford (1984), at Monash University (1984-85 and 1989-94), and at Princeton University (1985-89), before spending almost a decade at the Research School of Social Sciences at the Australian National University (1995-2004). He returned to Princeton University in 2004, where he is the McCosh Professor of Philosophy.

Smith's most influential work is *The Moral Problem* (1994), but he has also authored or co-authored more than ninety papers, a selection of which were published in a second book, *Ethics and the A Priori* (2004). A number of the papers he co-authored with Frank Jackson and Philip Pettit were collected together in *Mind, Morality and Explanation* (2004). Smith has also edited or co-edited four volumes. In addition to his well-known work on metaethics and moral psychology, he has made important contributions to normative ethics, political philosophy, philosophy of mind, and philosophy of action. His contributions to philosophy led to him being awarded an American Philosophical Association book prize in 2001 (for *The Moral Problem*), and a Centenary Medal for services to Australian society and the Humanities in 2003. Earlier, he was elected Fellow of the Academy of Humanities in Australia in 1997, and Fellow of the Academy of Social Sciences in 2000. It is safe to say that Smith is one of the most important moral philosophers of the late twentieth and early twenty-first centuries.

The problem that Smith confronts in *The Moral Problem* is the puzzle of how to reconcile moral realism with two theses that together provide us with a highly attractive story about how moral motivation, moral judgment and action are related to each other, and together seem to be inconsistent with the cognitivism required for moral realism (moral realists take moral judgments to be beliefs that are sometimes true). The first thesis, which goes by the name of *internalism*, states that a person who judges that it is right to do a particular act will, *ceteris paribus*, be motivated to do that act. This idea can be traced back to David Hume, but it was given special

prominence in the work of two twentieth-century Oxford philosophers, J. L. Mackie and Bernard Williams, who took it to pose a serious challenge to moral realism. Mackie argued that it is essential to the moral enterprise that we take there to be properties in the world that motivate us when we come into contact with them, and that a commitment to such 'queer' properties is not consistent with a naturalistic worldview (hence, morality involves a fundamental error). Williams argued that internalism places constraints on the practical reasons individuals can be thought to possess for acting in one way or another, and that putative moral demands cannot provide practical reasons for action for any individuals that do not already possess psychological states that could lead them to be motivated to act on such demands after a process of rational reflection (where this process, while it might lead to the creation of new desires, is understood to always be constrained by the desires that the individual already possesses).

The second thesis that Smith focuses on when setting out his main problem concerns motivation. This thesis, the *Humean theory of motivation*, consists of a sharp distinction between the belief states and desire states that are thought of as together explaining human actions (a desire is an 'original existence', to echo Hume on the passions), in combination with a contention that while desires play a crucial direct role in human behaviour, beliefs alone are not capable of moving us to act. Belief is in the business of merely representing the world, rather than attempting to change it, whereas desire contains no 'representative quality' (to echo Hume again; see *Treatise* 2.3.3).

Cognitivism, internalism, and the Humean theory of motivation seem to form an inconsistent triad, since cognitivism states that a person who judges that it is right to do a particular act *believes* that it is right to do that act, and internalism seems to imply that such a person will be motivated to do that act, but the Humean theory of motivation denies that there is any necessary connection between believing that it is right to do a particular act and being motivated to do it. In other words, internalism and the Humean theory of motivation seem to together imply that non-cognitivism is true, since if it is both true that moral judgments motivate us and that only desires motivate us, moral judgments cannot be beliefs.

Something has dropped out of this little argument for seeing the three main claims as inconsistent (as I have just presented it), and that is the '*ceteris paribus*' clause that must be included in the definition of internalism in order to ensure that the thesis is attractive, instead of implausible. Without such a clause, internalism would rule out the possibility of *akrasia*: judging that it is right to do a particular act, yet not being motivated to do it. But it is clear that people do suffer from *akrasia* and other forms of practical irrationality that prevent them from being

motivated to act on their best judgments. Although Smith initially states the internalist thesis by using a '*ceteris paribus*' clause, he goes on to argue that internalism can be understood, more specifically, as a conceptual necessity claim of the form, "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" (1994: 61). In other words, all things are equal just when an individual is being practically rational.

It is important to note that Smith doesn't think that merely recognising that the most plausible form of internalism is a weaker thesis than one might have initially supposed when setting out the main problem (ignoring the *ceteris paribus* clause, as I did above when explaining why the three main claims seem to be inconsistent) will enable us to avoid his trilemma. In a symposium on Smith's book published in *Ethics*, David Brink contends that this is all that is needed to avoid the trilemma, since only a strong internalist claim is inconsistent with the conjunction of cognitivism and the Humean theory of motivation (Brink 1997). Smith rejects this claim in his symposium response (Smith 1997), but even if Brink is right about what is required for there to be strict inconsistency here, Smith's project can still be appreciated for the way in which it attempts to give an account of how, contrary to appearances, the three main claims harmoniously cohere with each other (as well as for many other reasons).

The solution Smith provides to his main problem has many parts, but, in essence, it takes the following form. We can accept that belief and desire are distinct existences, and that no belief can directly lead to action, while rejecting Hume's contention that desires are not rationally criticizable and thus not answerable to beliefs. In particular, we can form beliefs about what our idealized selves would desire that our actual, non-ideal selves do in the circumstances we find ourselves in, where our idealized selves posses fully rational beliefs and desires (so far as their desires are concerned, this means that they have a set of desires that are maximally unified and coherent). In forming such beliefs we are, in effect, forming beliefs about what we have most reason to do. We are also forming beliefs that can play a special role in shaping our actual desires, since it is necessarily practically irrational for me to not want to do what I judge I would want myself now to do if I were only more rational. This last claim is particularly interesting and original. Beliefs about what a fully rational version of myself would desire my present rationally imperfect self to do seem to be just the right kinds of beliefs to play a strong normative role in relation to my desires.

Apart from providing a solution to the problem of reconciling cognitivism with internalism and the Humean theory of motivation, *The Moral Problem* also provides an account of the moral facts that make moral beliefs true (when they are true). Smith argues that when I imagine becoming more rational and more informed I must *perforce* also think of myself as converging with others in my fundamental desires. It is this convergence that assures us of the objectivity of morality. Facts about our normative reasons are facts about what we would all want our non-ideal selves to do if we were fully rational and fully informed, and moral facts are facts about what we would all desire that are of the appropriate substantive kind (where our understanding of this kind is fixed by moral platitudes that are required for possessing the very concept of 'right' to begin with).

Mackie's concern that moral facts (if there are any) must be very 'queer' since they motivate us whenever we encounter them is squarely dealt with, since to encounter a moral fact (as a moral fact) is just to encounter the sort of thing that it now naturally makes sense to think of as making it rational to desire to act in one way or another, i.e. a fact about what our more rational selves would want us to do. Williams' contention that moral demands will not necessarily provide every person with reasons to act (because some of us may not have desires that such demands can ever latch onto) can be rejected, without rejecting the spirit of Williams' defense of internalism, because moral demands just are the demands (of the appropriate substantive kind) that all rational people will converge on.

In his work after *The Moral Problem*, Smith focuses on a wide range of different problems, some of them fairly unconnected to the concerns of his first book. For example, the topics of freedom, responsibility and self-control are explored in the essays in *Ethics and the A Priori*, alongside essays closer in their concerns to his earlier work (essays that are concerned with moral realism and its alternatives, the nature of moral judgments, and the nature and demands of practical rationality). When it comes to normative ethics, Smith has elsewhere developed and argued for a type of actual-value consequentialism that incorporates the agent-relative goods traditionally thought to be incompatible with consequentialism.

In his metaethics, Smith's more recent work is somewhat continuous with the position mapped out in *The Moral Problem*. Two closely related ways in which he has moved away from the first book – possibly due in part to responses from critics – are that he is now less sanguine about the possibility of convergence in fundamental desires under conditions of increasing rationality, and he has developed a more substantive and dynamic conception of rationality. Where he earlier

sees it as a fairly easily recognizable conceptual truth that we would all converge in our fundamental desires were we to become fully rational, Smith has come to view this issue as presenting us with something more akin to an open question. He continues to hold that either we would converge, or there are no moral facts (externalists, on the other hand, would contend that there is a third possibility here), but his later work can be viewed as taking the threat of nihilism more seriously than his earlier work does.

It is particularly tempting to think we would not in fact converge in our desires were we to become fully rational as long as one sticks to a fairly thin notion of what rationality requires in relation to our desires – i.e., as long as one sticks to the original idea that what we are after is simply a maximally coherent and unified desire set (although it should be said that Smith was already adamant in his earlier work that it is not possible to provide a reduction of the concept of rationality to other concepts). Smith argues in "Internalism's Wheel", reproduced in *Ethics and the A Priori*, that we should think of our understanding of rationality as itself open to rational revision in a way that is very much susceptible to being influenced by our substantive judgments concerning our normative reasons and the demands of morality (he admits that there is thus a degree of circularity in his theory, but he contends that this circularity is virtuous, rather than vicious). If the later Smith is right that philosophical accounts of rationality and morality cannot be developed independently of each other, then the claim that we would all converge in our fundamental desires under conditions of full rationality – a claim whose truth Smith still believes to be necessary for there to be any moral facts – is an attractive claim indeed.

References

Brink, D. (1997) "Moral Motivation", *Ethics* 108: 4-32.

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

Smith, M. (1997) "In Defense of *The Moral Problem*: A Reply to Brink, Copp, and Sayre-McCord," *Ethics* 108: 84-119.

Smith, M. (2004) *Ethics and the A Priori: Selected Essays on Moral Psychology and Meta-ethics* New York: Cambridge University Press.

Jackson, F., Pettit, P. and Smith, M. (2004) *Mind, Morality, and Explanation: Selected Collaborations* Oxford: Clarendon Press.