

MICHAEL SMITH

Desire-based reason

Q1: why should I care about the rationality of the moral agent? what about emotion?

MICHAEL SMITH (1954-) is McCosh Professor of Philosophy at Princeton University. In this excerpt from his 1994 book *The Moral Problem*, Smith aims to provide an account of moral rightness that can be objective and have motivational force. An act is morally right, according to Smith, just in case it is the act that we would all want to do if we were fully rational. After presenting his analysis, Smith defends the existence of moral facts against the objections of Philippa Foot, Gilbert Harman, and J. L. Mackie.

Neo-Kantian

Issue: What do objective facts have to do with desire?

THE SOLUTION TO THE MORAL PROBLEM

We now have a solution to the moral problem at hand. That is, we can now explain why the following three propositions:

1. Moral judgements of the form "It is right that I Φ " express a subject's beliefs about an objective matter of fact, a fact about what it is right for her to do.
2. If someone judges that it is right that she Φ s then, *ceteris paribus*, she is motivated to Φ .
3. An agent is motivated to act in a certain way just in case she has an appropriate desire and a means-end belief, where belief and desire are, in Hume's terms, distinct existences.

are both consistent and true. Let me briefly explain why.

If our concept of rightness is the concept of what we would desire ourselves to do if we were fully rational, where this is a desire for something of the appropriate substantive kind, then it does indeed follow that our moral judgements are expressions of our beliefs about an objective matter of fact. For our moral judgements are expressions of our beliefs about what we have normative reason to do, where such reasons are

Michael Smith, *The Moral Problem* (Wiley-Blackwell, 1994).

is readily accepted, just because the motives for their objectification are less compelling.

But it would be misleading to think of the objectification of moral values as primarily the projection of feelings, as in the pathetic fallacy. More important are wants and demands. As Hobbes says, "whatsoever is the object of any man's Appetite or Desire, that is it, which he for his part calleth *Good*"; and certainly both the adjective "good" and the noun "goods" are used in non-moral contexts of things because they are such as to satisfy desires. We get the notion of something's being objectively good, or having intrinsic value, by reversing the direction of dependence here, by making the desire depend upon the goodness, instead of the goodness on the desire. And this is aided by the fact that the desired thing will indeed have features that make it desired, that enable it to arouse a desire or that make it such as to satisfy some desire that is already there. It is fairly easy to confuse the way in which a thing's desirability is indeed objective with its having in our sense objective value. The fact that the word "good" serves as one of our main moral terms is a trace of this pattern of objectification.

possible vs objective

Study Questions

1. Why does Mackie think objective moral values would have to be so queer? What features of objective values, according to Mackie, makes them unlike other properties or entities?
2. Mackie is aware that the mere fact of disagreement does not undermine the objectivity of the disputed issues. How does Mackie distinguish moral disagreement from scientific disagreement, an area that does admit of objective treatment?
3. Mackie and Ayer agree that there are no objective moral values. Which of the two antirealist positions do you find more plausible and why?

Conceptual claim:

Rightness in circumstances C is the feature we would want acts to have in C if we were fully rational, where these wants have the appropriate content

Substantive claim:

Fness is the feature we would want acts to have in C if we were fully rational, and Fness is a feature of the appropriate kind

Conclusion:

Rightness in C is Fness

And this argument is, in turn, broadly naturalistic in two respects. First, it is naturalistic in so far as the features that we would want our acts to have under conditions of full rationality, the features that we would want acts to instantiate in the evaluated possible world, are themselves all natural features whenever the evaluated world is itself naturalistic. Our non-reductive, summary style definition of rightness, in conjunction with a substantive claim of the kind described, thus allows us to identify rightness with a natural feature of acts in naturalistic worlds like the actual world: for example, in this case, with Fness.

And second, even though the analysis is not itself naturalistic—even though it defines rightness in terms of full rationality where this may not itself be definable in naturalistic terms—fully rational creatures in the evaluating possible world are themselves naturalistically realized. For a fully rational creature is simply someone with a certain psychology and, as you will recall, a natural feature is simply a feature that figures in one of the natural or social sciences, *including psychology* (chapter 2). Of course, the psychology of a fully rational creature is an idealized psychology, but such an idealization requires nothing non-natural for its realization. Thus, if we wanted to, we could construct non-reductive analyses of the key normative concepts we use to characterize the normative features of such

an idealized creature's psychology—the unity, the coherence, and the like, of its desires—and then use these analyses to construct two-stage arguments, much like that just given, in order to identify these normative features of a fully rational creature's psychology with natural features of its psychology (for an analogy, see note 9 to chapter 2). Coherence and unity, though not naturalistically definable are therefore themselves just natural features of a psychology. The evaluating possible world is therefore naturalistic in the relevant respect as well.

The analysis of rightness provided thus makes the legitimacy of moral talk depend ultimately upon the possibility of identifying moral features, like rightness, with natural features of acts. Absent such identifications, we would have to conclude that moral features are simply not instantiated at all; that moral talk is, much as Mackie thought, based on an error of presupposition. But can we say more? Can we say whether moral talk is or is not legitimate?

ARE THERE ANY MORAL FACTS?

The substantive claim in the two-stage argument described above tells us that moral talk is legitimate just in case a certain condition is met. My handing back a wallet I found in the street in such and such circumstances is right, for example, only if, under conditions of full rationality, we would all want that if we find a wallet in the street in such and such circumstances, then we hand it back. Of course, if this is indeed true, then it is an *a priori* truth. The fact that we would have such a desire under conditions of full rationality will be a consequence of the theory that systematically justifies our desires, and this task of theory construction is itself a relatively *a priori* enterprise; it is a task that requires reflection and conversation, not empirical investigation. However that does not mean that it is an *obvious* truth. For it might not only take a good deal of reflection and conversation for any

an idealized creature's psychology—the unity, the coherence, and the like, of its desires—and then use these analyses to construct two-stage arguments, much like that just given, in order to identify these normative features of a fully rational creature's psychology with natural features of its psychology (for an analogy, see note 9 to chapter 2). Coherence and unity, though not naturalistically definable are therefore themselves just natural features of a psychology. The evaluating possible world is therefore naturalistic in the relevant respect as well.

The analysis of rightness provided thus makes the legitimacy of moral talk depend ultimately upon the possibility of identifying moral features, like rightness, with natural features of acts. Absent such identifications, we would have to conclude that moral features are simply not instantiated at all; that moral talk is, much as Mackie thought, based on an error of presupposition. But can we say more? Can we say whether moral talk is or is not legitimate?

ARE THERE ANY MORAL FACTS?

The substantive claim in the two-stage argument described above tells us that moral talk is legitimate just in case a certain condition is met. My handing back a wallet I found in the street in such and such circumstances is right, for example, only if, under conditions of full rationality, we would all want that if we find a wallet in the street in such and such circumstances, then we hand it back. Of course, if this is indeed true, then it is an *a priori* truth. The fact that we would have such a desire under conditions of full rationality will be a consequence of the theory that systematically justifies our desires, and this task of theory construction is itself a relationally *a priori* enterprise; it is a task that requires reflection and conversation, not empirical investigation. However that does not mean that it is an *obvious* truth. For it might not only take a good deal of reflection and conversation for any

individual to discover this to be true, it might also take time and effort to convince anyone else. And of course, even after all of that time and effort, it might turn out that we are wrong. What we thought was an *a priori* truth might have been no truth at all. In deciding whether or not moral talk is legitimate, then, it seems to me that we have no alternative but to admit that we are venturing an opinion on something about which we can have no cast-iron guarantee.

However, for all that, it seems to me that we should none the less have some confidence in the legitimacy of moral talk. For, in short, the empirical fact that moral argument tends to elicit the agreement of our fellows gives us reason to believe that there will be a convergence in our desires under conditions of full rationality. For the best explanation of that tendency is our convergence upon a set of extremely unobvious *a priori* moral truths. And the truth of these unobvious *a priori* moral truths requires, in turn, a convergence in the desires that fully rational creatures would have.

Now this argument is likely to meet with some resistance. After all, isn't there currently much entrenched moral disagreement? And don't such disagreements constitute profound obstacles to a convergence in our desires in fact emerging? This is true, but it does not count against the force of the argument. Indeed, once we remember the following three points, such disagreements can be seen to add to the force of the argument.

First, we must remember that alongside such entrenched disagreements as we in fact find we also find massive areas of entrenched agreement. As I see it, this is the real significance of the fact that we have and use the so-called "thick" moral concepts, concepts that at once both describe some naturalistic state of affairs and positively or negatively evaluate it: concepts like courage, brutality, honesty, duplicity, loyalty, meanness, kindness, treachery, and the like (Williams, 1985: 129). For what the prevalence of such concepts suggests is that there is in fact considerable agreement about what is right and

wrong: acts of brutality, duplicity, meanness and treachery are wrong, at least other things being equal, whereas acts of courage, honesty, loyalty, and kindness are right, again, other things being equal. What the prevalence of such concepts suggests is therefore that moral agreement is in fact so extensive that our language has developed in such a way as to build an evaluative component into certain naturalistically descriptive concepts.

Second, when we look at current areas of entrenched disagreement, we must remember that in the past similarly entrenched disagreements were removed *inter alia* via a process of moral argument. I am thinking in particular of the historical, and in some places still current, debates over slavery, worker's rights, women's rights, democracy and the like. We must not forget that there has been considerable moral progress, and that what moral progress consists in is the removal of entrenched disagreements of just the kind that we currently face.

And third and finally, we must remember that where entrenched disagreements currently seem utterly intractable we can often explain why this is the case in ways that make them seem less threatening to the idea of a convergence in the opinions of fully rational creatures. For example, one or the other parties to the disagreement all too often forms their moral beliefs in response to the directives of a religious authority rather than as the result of the exercise of their own free thought in concert with their fellows. But beliefs formed exclusively in this way have dubious rational credentials. They require that we privilege one group's opinions about what is to be done—those of the religious authority—over another's—those of the followers—for no good reason. The fact that disagreement persists for this sort of reason thus casts no doubt on the possibility of an agreement if we were to engage in free and rational debate.

In light of these points it seems to me that, notwithstanding such disagreements as there are and will perhaps remain, we should therefore in fact be quite optimistic about the possibility

of an agreement about what is right and wrong being reached under more idealized conditions of reflection and discussion. We might eventually come to be pessimistic, of course. Our epistemic situation might deteriorate, widespread disagreements might emerge, disagreements that seem both unresolvable and inexplicable. And if that were to happen, then we might well quite justifiably come to think that Mackie was right after all, that there are no moral facts, though there would still be room for doubt. The point is simply that this *is not* our current epistemic situation. [...]

Foot's and Harman's Objection

Foot's main objection to rationalism is that we can give no content to the idea that the immoral person is irrational.

Attempts have sometimes been made to show that some kind of irrationality is involved in ignoring the "should" of morality: in saying "Immoral—so what?" as one says "Not *comme il faut*—so what?" But as far as I can see these have rested on some illegitimate assumption, as, for instance, of thinking that the amoral man who agrees that some piece of conduct is immoral but takes no notice of that, is inconsistently disregarding a rule of conduct that he has accepted; or again of thinking it inconsistent to desire that others will not do to one what one proposes to do to them. The fact is that the man who rejects morality because he sees no reason to obey its rules can be convicted of villainy but not inconsistency. (1972: 161)

he begins: that he has a normative reason to gain wealth no matter what the cost to others. For, as we have seen, this is equivalent to the claim that fully rational creatures would want that, if they find themselves in the circumstances of the successful criminal, then they gain wealth no matter what the cost to others. And the successful criminal's opinion notwithstanding, it seems quite evident that we have no reason to believe that this is true. Fully rational creatures would want no such thing.

Note what I have not said. I have not said that the fact that we all disagree with the successful criminal entails that he is wrong. Perhaps we are all mistaken about what fully rational creatures would want. But the mere fact that it is logically possible that we are wrong gives us no more reason to endorse the opinions of the successful criminal and doubt our own convictions than the mere fact that it is logically possible that we are wrong when we think that the sun will rise tomorrow gives us reason to endorse the opinions of the prophets of doom.

Harman tells us that the fact that the successful criminal does not observe the moral requirement not to harm others cannot be traced to a "failure to consider or appreciate certain arguments." But it seems to me that that should now seem quite wrong. After all, the successful criminal thinks that he has a normative reason to gain wealth no matter what the cost to others, and he sticks with this opinion despite the fact that virtually everyone disagrees with him.

Moreover, he does so without good reason. For he can give no account of why his own opinion about what fully rational creatures would want should be privileged over the opinion of others; he can give no account of why his opinion should be right, others' opinions should be wrong. He can give no such account because he rejects the very idea that the folk possess between them a stock of wisdom about such matters against which each person's opinions should be tested. And yet, ultimately, this is the only court of appeal there is for claims about what we have

appreciate certain arguments, ignorance of relevant evidence, errors in reasoning, irrationality, unreasonableness, or weakness of will...

The absolutist might argue that the criminal must be irrational or at least unreasonable. Seeing that a proposed course of action will probably cause serious injury to some outsider, the criminal does not treat this as a reason not to undertake that course of action. This must be irrational or unreasonable, because such a consideration simply is such a reason and indeed is an obvious reason, a basic reason, not one that has to be derived in some complex way through arcane reasoning. But then it must be irrational or at least unreasonable for the criminal not to care sufficiently about others, since the criminal's lack of concern for others is what is responsible for the criminal's not taking the likelihood of harm to an outsider to be a reason against a proposed course of action...

The relativist's reply to such an argument is that, on any plausible characterization of reasonableness and unreasonableness (or rationality and irrationality) as notions that can be part of the scientific conception of the world, the absolutist's claim is just false. Someone can be completely rational without feeling concern and respect for outsiders. (1985: 39-40)

In effect, then, both Harman and Foot challenge the rationalist to explain an uncontroversial sense in which the successful criminal is being irrational. Contrary to both, however, it seems to me easy to meet this challenge.

The successful criminal must begin his deliberations from some evaluative premise or other. Let's imagine that he begins from the premise that he has a normative reason to gain wealth no matter what the cost to others. Now, given this premise we can certainly imagine that he is not inconsistent, that he does not disregard a rule of conduct that he has himself already accepted, that he is not inattentive, that he does not suffer from weakness of will and so on and so forth... But all that that shows is that the flaw in his reasoning lies in the premise from which

normative reason to do. The successful criminal thus seems to me to suffer from the all too common vice of *intellectual arrogance*. He therefore does indeed suffer from a "failure to consider or appreciate certain arguments," for he doesn't feel the force of arguments that come from *others* at all. [...]

Mackie's Objections

John Mackie tells us that

... Kant himself thought that moral judgements are categorical imperatives... So far as ethics is concerned, my thesis that there are no objective values is specifically the denial that any such categorically imperative element is objectively valid. (1977: 29)

Mackie's two famous objections to the claim that there exist objective values are thus directed quite specifically against the rationalist. Consider his first argument.

If there were objective values, then they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe. Correspondingly, if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from ordinary ways of knowing everything else. (1977: 38)

But can Mackie really lay a charge of strangeness against rightness, at least as that feature of acts has been analysed here? It seems not.

To say that performing an act of a certain sort in certain circumstances is right is, I have argued, to say *inter alia* that there is a normative reason to perform it. And this, in turn, is simply to say that fully rational creatures would desire that such an act be performed in such circumstances, where such a desire is of the appropriate substantive kind. On this account, moral features like rightness thus simply are not "entities or qualities or relations of a very strange sort"; no "special faculty or moral perception of intuition" is required in order to gain knowledge of them. All that is required is the ability to think about

what a more rational person would want. And this, in turn, just requires the ability to engage in an attempt to find a systematic and unified justification of the various desires that we have; to engage in a process of reflective equilibrium. The charge of strangeness is thus entirely misplaced.

Consider now Mackie's second argument. The idea here is that

...radical differences between first order moral judgements make it difficult to treat those judgements as apprehensions of objective truths. But it is not the mere occurrence of disagreements that tells against the objectivity of values. Disagreement on questions in history or biology or cosmology does not show that there are no objective issues in these fields for investigators to disagree about. But such scientific disagreement results from speculative inferences or explanatory hypotheses based on inadequate evidence, and it is hardly plausible to interpret moral disagreement in the same way. Disagreement about moral codes seems to reflect people's adherence to and participation in different ways of life. (1977: 36)

Mackie thus agrees that the availability of a rational procedure for resolving moral disagreement would render the "mere occurrence" of moral disagreements harmless, from the rationalist's point of view. But he suggests that the idea that there exists such a procedure is "hardly plausible." His actual arguments for this claim are, however, hardly convincing.

Mackie tells us that moral disagreement cannot plausibly be interpreted as resulting from "speculative inferences or explanatory hypotheses based on inadequate evidence." But, in many ways, this is just false. For, as we have seen, when we construct moral theories we are trying to find out what we have normative reason to do, and when we try to find out what we have normative reason to do we are trying to find out what fully rational creatures would want, and when we try to find out what fully rational creatures would want we embark on a procedure of justification of our various desires that is very similar indeed

to the enterprise of theory construction in science in response to observational evidence. Both are simply applications of the idea of reflective equilibrium.

Of course, perhaps Mackie is saying that this attempt to provide our desires with a systematic justification is doomed to failure; that no such justification will be forthcoming; that there is no single thing that fully rational creatures would all want us to do in the various circumstances in which we find ourselves. Perhaps that is the real point of his argument from relativity. But if so then the reply is again simple enough. For though it is certainly logically possible that no systematic justification of our desires is forthcoming, in light of the remarks we made in the previous section about the power of moral argument to elicit agreement, it seems more reasonable to think that such a justification is forthcoming.

This is not to deny, of course, that the real test of Mackie's argument from relativity lies in the ultimate outcome of debate in normative ethics. The real question is whether we will, by engaging in such debate, come up with answers to moral questions that secure the free agreement of those who participate. If we do not, then our confidence that such agreement is to be had may be undermined, and quite rightly so. But that gives us no reason for scepticism now. We must give the arguments and see what their outcome is.

References

- Foot, Philippa. 1972. "Morality as a System of Hypothetical Imperatives," reprinted in Foot (1978), pp. 157-173.
- . 1978. *Virtues and Vices*. Berkeley, CA: University of California Press.
- Harman, Gilbert. 1977. *The Nature of Morality*. Oxford: Oxford University Press.
- Mackie, John. 1977. *Ethics: Inventing Right and Wrong*. London: Penguin.
- Williams, Bernard. 1985. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

Def re flechm
equilibrium

Source
Mackie's
argument