Chapter 5

Moral Intuition

Jeff McMahan

Moral Inquiry

Suppose we wish to understand a particular moral problem – for example, abortion. One way of proceeding is to reason on the basis of our existing substantive moral beliefs. We may, however, suspect that our moral beliefs about abortion, insofar as we have any prior to serious reflection, are unreliable. We may suspect, for example, that our beliefs about abortion reflect the influence of a religious education that we now repudiate, or that our feminist sympathies may make us insufficiently sensitive to the status of the fetus. Thus we may take as our starting point certain related moral beliefs about which we are more confident – for example, that killing an unthreatening, morally innocent, adult human being (or "person," for the sake of brevity) whose continued life would be worth living is, except perhaps in the most extreme circumstances, seriously wrong, while painlessly killing a lower nonhuman animal (for example, a mouse) may be permissible provided that the interests that are thereby served outweigh those of the animal that would be frustrated by its death. There is, of course, divergence of opinion even about these cases. Some people believe that intentionally killing a person can never be justified, while others believe that it can be justified whenever it is necessary to save the lives of a greater number of innocent people. And, while some believe that there is no objection whatever to killing a mouse independently of the effects this might have on human interests, others believe that killing a mouse is seriously objectionable just because of the effect on the mouse and requires a strong justification in order to be permissible. Nevertheless, everyone agrees that, in ordinary circumstances, killing a person is more objectionable morally than killing a mouse.

We could therefore initiate our inquiry into abortion by exploring our confident sense that there is a significant moral difference between killing a person and killing a lower animal – so that, for example, the killing of a lower animal might be justified by appeal to considerations that would not constitute even the beginning of a justification for killing a person. With these beliefs as our starting point, we could work our way toward a better understanding of abortion. We could proceed by trying to understand why killing a person is generally wrong and why it is generally so much more seriously wrong than killing a lower animal. What are the relevant differences between a person and a lower animal? Are the properties of persons that make killing them generally worse all intrinsic properties? Or is part of the explanation of the greater wrongness of killing persons that we normally bear certain relations to them that do not exist between ourselves and animals? In addressing these questions, we may consult our intuitions about a range of particular cases and this may yield provocative results. We may notice, for example, that the extent to which killing an animal seems wrong varies with the degree of harm the animal suffers in dying. Thus it seems more objectionable to kill a dog than to kill a mouse; and the explanation seems connected with the fact that the dog loses more by dying. But we may also notice that the extent to which it is wrong to kill a person does not seem to vary with the extent to which death is bad for the victim. Thus it seems no less wrong, other things being equal, to kill a dullard than to kill a genius, or to kill an elderly person with a reduced life expectancy than to kill a person in the prime of life.

As our understanding of the morality of killing in general increases, we can begin to extract from our findings various implications for the morality of abortion. Suppose, for example, that we discover that there are certain properties that adult human beings generally possess that lower animals do not that seem to help explain the difference between killing people and killing animals. We can then consider whether these properties are possessed by human fetuses. If they are, then in that respect abortion is relevantly like the killing of a person; if not, there is then reason to suppose that abortion is morally more like the killing of animals.

These remarks about abortion are intended only to provide a sketchy illustration of a certain approach to practical ethics, a certain general pattern of reasoning about moral problems. Its most conspicuous feature is that it treats certain substantive moral beliefs that we already have as prima facie reliable starting points for moral inquiry. It presupposes that at least some of our moral intuitions have a certain presumptive epistemic authority.

Intuitions

What are moral intuitions? As I will understand the term, a moral intuition is a moral judgment – typically about a particular problem, a particular act, or a particular agent, though possibly also about a moral rule or principle – that is not the

result of inferential reasoning. It is not inferred from one's other beliefs but arises on its own. If I consider the act of torturing the cat, I judge immediately that, in the circumstances, this would be wrong. I do not need to consult my other beliefs in order to arrive at this judgment. This is not to say that a moral intuition is necessarily elicited instantaneously, the way a sense perception is. If a particular problem or case is complex, one may have to consider it at length in order to distinguish and assimilate its various relevant features – in much the same way that one might have to examine the many details of a highly complex work of art in order to judge or appreciate it.

The belief that I cited as one of the possible starting points for an inquiry about abortion – namely, that killing a person is generally wrong – may or may not count as an intuition according to this understanding, though for most of us it has an intuitive basis. It is an intuition if one finds it immediately compelling, but not if one accepts it as an inductive inference from one's intuitively finding that in this, that, and the other case, killing a person is wrong, or if one deduces it from the principle that whatever God prohibits is wrong.

In the history of moral philosophy, the idea that moral intuitions have epistemic authority has been associated, unsurprisingly, with a cluster of theories that have traveled under the label "intuitionism." Those doctrines are many and various and I do not propose to disentangle them. But two claims associated with certain historically prominent variants of intuitionism have done much to discredit the appeal to intuitions. One is that intuitions are the deliverances of a special organ or faculty of moral perception, typically understood as something like an inner eye that provides occult access to a noumenal realm of objective values. The other is often regarded as a corollary of the first – namely, that intuitions are indubitable (that is, that their veridicality cannot be doubted) as well as infallible (that is, that they cannot in fact be mistaken). But a variety of considerations – such as the diversity of moral intuitions, the fact that people do often doubt and even repudiate certain of their intuitions, and the evident origin of some intuitions in social prejudice or self-interest – make it untenable to suppose that intuitions are infallible apprehensions of moral reality by some special faculty of moral perception.

There are other features that are occasionally attributed to intuitions that are in fact inessential. It is sometimes said, for example, that intuitions are "pretheoretical." If all this means is that they are not derived inferentially from a moral theory, then this of course follows from the stipulation that they are not the products of any sort of inferential reasoning. If, however, it means that intuitions must be untutored or entirely unaffected by a person's exposure to moral theory, then the requirement is evidently too strong. Just as many people's moral intuitions have been shaped by their early exposure to religious indoctrination, so some people's intuitions are gradually molded by their commitment to a particular moral theory. The stipulation that intuitions are not the products of conscious inference does not entail that they cannot be affected by learning. That they may arise spontaneously is compatible with their having sources in one's nature that are malleable.

O

Theory

Many philosophers reject the idea that moral intuitions have epistemic authority. Peter Singer, for example, has suggested that we should assume "that all the particular moral judgments we intuitively make are likely to derive from discarded religious systems, from warped views of sex and bodily functions, or from customs necessary for the survival of the group in social and economic circumstances that now lie in the distant past." On this assumption, he notes, "it would be best to forget all about our particular moral judgments" (Singer 1974: 516). It is, of course, possible to be less dismissive of intuitions and yet still regard them as lacking in authority. Some philosophers, for example, concede that intuitions may be reasonably reliable guides to action in most circumstances – since morality must ensure that people are equipped with dispositions to believe and act in certain ways in situations in which deliberation and reflection are not possible – but deny that they are a source of moral knowledge or have any proper role in reasoning about moral problems. They believe that practical reasoning about a moral problem must consist in determining what some favored moral theory implies about the problem. It is the theory that is the source of our moral knowledge concerning particular problems and cases. And the theory is itself validated by means other than its conformity with intuition.

According to this approach, if our concern is to understand the morality of abortion, our first task must be to discover the correct moral theory. Moral inquiry is initially and primarily theoretical; only at the end of this theoretical inquiry is it possible to address moral problems such as abortion competently, bringing the theory to bear and extracting from it the knowledge we initially sought. This general approach therefore contrasts with the first approach I sketched, according to which moral inquiry begins with problems and cases and our intuitions about them, seeks principles that unify and explain the intuitions, and proceeds through adjustment and modification of both the principles and intuitions until consistency and harmony are achieved. On this approach, a moral theory in which we are entitled to have confidence is something that we can hope to have only near the end of the process of inquiry into problems of substantive morality.

Let us refer to the two broadly defined patterns of moral inquiry that I have sketched as the *Intuitive Approach* and the *Theoretical Approach*. Both are richly represented in the history of moral philosophy. The Socrates of Plato's dialogues is an admirable exemplar of the Intuitive Approach, while Hobbes and Kant exemplify the Theoretical Approach. Each of the latter begins with a conception of the nature of morality that he believes dictates a particular method for arriving at moral judgments about particular problems and cases. In recent years, most philosophers working on problems of practical ethics have largely followed the Intuitive Approach, but the Theoretical Approach also has many distinguished recent exponents, among them Richard Hare, Richard Brandt, and an assortment of theorists in the contractualist and consequentialist traditions.

The Theoretical Approach is reformist in a rather radical way. People have always reasoned and argued about substantive issues in morality. According to adherents of the Theoretical Approach, however, people have been misguided insofar as their reasoning has diverged from the forms and patterns of moral reasoning prescribed by the correct moral theory. Richard Brandt, for example, suggests that "is morally wrong" means "would be prohibited by any moral code which all fully rational persons would tend to support, in preference to all others or none at all, for the society of the agent, if they expected to spend a lifetime in that society" (Brandt 1979: 194). Assuming that this definition also states a test for determining whether an act is wrong, it seems clear that any convergence of the conclusions of most people's actual moral reasoning and the conclusions that might result from Brandt's proposed mode of reasoning would be to a considerable degree fortuitous or coincidental. According to the Theoretical Approach, therefore, philosophical ethics is utterly different from, say, the philosophy of science. While the philosopher of science may criticize certain aspects of the practice of science, and may urge scientists to revise their understanding of the nature of their practices or the status of their conclusions, the philosopher does not presume to tell scientists that they have been utilizing the wrong method and would do better to adopt a different approach. The Intuitive Approach is in general more respectful of the modes of moral reasoning that people actually employ – though only because people in fact tend to reason about moral problems in the way it recommends.

Theory Unchecked by Intuition

Could we really conduct our thinking about morality and moral problems in the way suggested by the Theoretical Approach, without building up from our moral intuitions or consulting those intuitions to test the plausibility of the implications of proposed moral theories? Even those who most vehemently deny that intuitions have any independent credibility nevertheless often build their arguments on the basis of appeals to common intuitions (for example, Rachels 1986: 112–13, 134–5; Singer 1993: 229). But, although this is suggestive of the difficulty of producing arguments capable of persuading people without linking them to our preexisting moral beliefs, it is merely an ad hominem point and as such does little to support the appeal to intuition. An alternative point that may be urged against the Theoretical Approach is that our intuitions often *compel* belief in a way that, for most of us, no moral theory does. If an intuition that is highly compelling cannot be reconciled with what seems to be the best supported moral theory, can it be rational to reject the intuition at the behest of a theory that is feebler in its ability to compel belief?

It is important to be clear about the nature of this challenge. The claim is not simply that moral intuitions often strike us as more obvious or less open to

doubt than it seems any moral theory is. By itself, this would not be a strong consideration in favor of the intuition. The theories of modern physics tell us that many of our commonsense beliefs about the nature of the physical world are mistaken. Many of these beliefs seem overwhelmingly obvious while the theory that disputes them may be so arcane as to be unintelligible to all but a few. Yet most of us recognize that some scientific theories that overturn aspects of our commonsense conception of the physical world are so well established by their powers of explanation and prediction and by the control they give us over the forces of nature that we readily acquiesce in their claims and concede that our commonsense views must be illusory. If a moral theory could command our allegiance by comparable means of persuasion, we might yield our intuitions to it without demur, even if it had none of the immediate obviousness in which our intuitions tend to come clothed. But the challenge to the Theoretical Approach is that no moral theory, at least at the present stage of the history of philosophical ethics, can have the degree of authority or validation that the best supported scientific theories have. The lamentable truth is that we are at present deeply uncertain even about what types of consideration support or justify a moral theory. There are no agreed criteria for determining whether or to what extent a moral theory is justified. So when an intuition, which may be immediately compelling, comes into conflict with a moral theory, which can have nothing approaching the authority of a well-grounded scientific theory, it is not surprising that we should often be reluctant to abandon the intuition at the bidding of the theory. We can, indeed, be reasonably confident *in advance* that none of the moral theories presently on offer is sufficiently credentialed to make it rationally required that we surrender our intuition.

It is instructive to consider how most of us respond when, on inquiring into a particular moral problem, we find that some moral theory has implications for the problem that clash with our intuitions. Our response is not to question how well grounded the theory is, on the assumption that we should be prepared to acquiesce if we find that the theory is well supported. If the theory generates its conclusion via a distinct argument, our tendency is to detach the argument from the parent theory and consider it on its own merits. According to R.M. Hare, for example, his universal prescriptivist theory of morality implies that we should reason about the morality of abortion by applying a variant of the Golden Rule: "We should do to others as we are glad that they did do to us" (Hare 1975: 208). When we discover that, at least according to Hare, this principle implies not only that abortion is generally wrong but also that remaining childless is as well, we do not go back to Hare's earlier books to check the arguments for universal prescriptivism. Instead we undertake an independent inquiry to try to determine whether and, if so, to what extent it matters to the morality of abortion that, when an abortion is not performed, there will typically later be a person who is glad to exist who would not have existed if the abortion had been performed. That is, if we are serious about understanding the morality of abortion, we will take seriously the considerations identified as relevant by the theory; and we may be grateful to

the theory for having helped us to see whatever relevance these considerations may in fact have; but we are generally not overawed by the fact that these considerations have been identified as relevant *by the theory*. Their provenance in the theory fails to imbue them with epistemic authority.

One might even wonder what claim a theory might have to be a *moral* theory if it has foundations that are wholly independent of the intuitions that have shaped the common features of all recognizably moral codes. Recall, for example, Brandt's claim that an act is morally wrong if it "would be prohibited by any moral code which all fully rational persons would tend to support, in preference to all others or none at all, for the society of the agent, if they expected to spend a lifetime in that society." It is possible that fully rational persons would tend to prefer the code that, if generally accepted, would make them (though not necessarily nonrational beings, such as animals) most comfortable. The methodology rules out the possibility that, in choosing such a code, they would be failing to recognize or respect their own moral status. (Recognition of their moral status could not, of course, be a condition of their being rational, for that would make the account circular.) I cite Brandt's theory for the purpose of illustration because it imposes fewer constraints on the choice of principles than most other contractualist and rule-consequentialist theories, and is therefore less likely than they are to yield principles that support common intuitions. But a similar concern arises for all theories that reject conformity with intuition as a constraint on moral justification.

Moral Epistemology

The remarks in the previous section are meant only to suggest certain reservations we might have about the Theoretical Approach; they are far from providing decisive reasons for rejecting that approach. Moreover, even if we had stronger grounds for skepticism about the Theoretical Approach than those I have offered, this would still be insufficient to compel us to accept the Intuitive Approach. For it is hardly a ground for confidence in our intuitions that there are reasons for doubting the approach to moral inquiry that denies them a role. Something more positive has to be said on behalf of our intuitions themselves. At a minimum, more needs to be said about the role they are supposed to have in the structure of justification in ethics. In what follows I will first offer a few general remarks about moral epistemology, after which I will briefly sketch an account of moral inquiry that explains the role that our intuitions might have in our moral thinking and also helps to elucidate their epistemological status. I will then conclude by noting that there is a conception of the nature of moral knowledge that has independent plausibility and, if correct, offers a deeper understanding of the epistemological status of moral intuitions.

A theory in epistemology may be a theory either of truth or of justification. I will focus on the issue of justification and simply assume that there is a tight

connection between justification and truth. Accounts of justification, including accounts of justified moral belief, tend to be divided into two major approaches: coherentism and foundationalism. Coherentist accounts of moral justification hold that a moral belief is justified solely in terms of its relations, particularly its inferential relations, with other beliefs. It is justified to the extent to which it coheres with a set of beliefs that together form a coherent whole. By contrast, foundationalist accounts hold that some beliefs are self-justifying – at least in the sense that they are justified independently of their relations to other beliefs. According to foundationalist accounts, a moral belief is justified if and only if it is either self-justifying or bears an appropriate inferential relation to a belief that is self-justifying.

Of the two types of account, coherentism is generally thought to be more hospitable to the Intuitive Approach. The most commonly endorsed method of moral inquiry among contemporary moral philosophers is the method described by John Rawls under the label "reflective equilibrium" (Rawls 1972: 19–21, 48–51). According to the method of reflective equilibrium, we begin with a set of moral intuitions about particular cases, filter out those that are the obvious products of distorting influences, and then seek to unify the remaining intuitions under a set of more general principles. We seek principles that both imply and explain our particular judgments. But the match between principles and intuitions will inevitably be very imperfect in the first instance. A candidate principle may imply a great many of our intuitions and yet have some implications that conflict with other intuitions. In that case we may modify or even abandon the principle; but, if the principle has considerable explanatory power with respect to a wide range of intuitions and cannot be modified without significant sacrifice of this power, we may instead decide to reject the recalcitrant intuitions. In this way we make reciprocal adjustments between intuitions and principles until our beliefs at various levels of generality are all brought into a state of harmony, or reflective equilibrium. This method is generally interpreted in coherentist terms, in that it is understood to make coherence with other beliefs the sole criterion of a belief's credibility. Yet it obviously treats intuitions as potential sources of moral knowledge. Although intuitions arise noninferentially and thus, in coherentist terms, have no prima facie credibility on their own, those that survive the initial filtration and are compatible with the principles that emerge in the process of trying to reach reflective equilibrium turn out to be justified moral beliefs.

Foundationalist theories of moral justification tend to be favored by proponents of the Theoretical Approach. Typically, the foundational beliefs (that is, those that are not justified in terms of their relations to other beliefs) are held to be nonmoral; justified moral beliefs are all ultimately derivable via some process of reasoning that is based on the foundational nonmoral beliefs (for examples, see Timmons 1987). Those who attribute authority to our moral intuitions tend, understandably, to be more reluctant to embrace foundationalism. This is mainly because it seems implausible to regard our intuitions themselves as foundational. This seems to attribute to them too exalted a status. While our intuitions do seem to have a

certain initial credibility, it seems exorbitant to suppose that they are self-evident or self-justifying. They seem at most to have an evidentiary status. We recoil from the suggestion (advanced, as I noted earlier, by various traditional intuitionists) that intuitions are the unshakable basis on which all moral knowledge rests.

There are, however, at least two ways of overcoming this ground of reluctance to combine foundationalism with the Intuitive Approach. The first is to recognize that a belief may be of the foundational *sort* and yet be defeasible. Suppose, for example, that sense perceptions are the foundations of empirical knowledge. Even if all empirical knowledge is derived immediately from sense perceptions or is ultimately traceable by chains of inference to sense perceptions, it does not follow that *all* sense perceptions are sources of empirical knowledge. Some may be distorted, illusory, or otherwise erroneous. And there is no reason why the same may not be true, *mutatis mutandis*, in the case of moral intuitions.

Second, a foundationalist account of moral knowledge may treat intuitions as reliable sources of moral knowledge without treating them as foundational or self-justifying. It is this possibility that I will explore in more detail.

A Sketch of a Foundationalist Conception of Moral Justification

To most moral philosophers who reason about substantive moral issues, it seems that the method of reflective equilibrium, or a process very similar to it, is the best or most fruitful method of moral inquiry. Of the known methods of inquiry, it is the one that seems most likely to lead to justified moral beliefs. It does not, however, have to be interpreted within the coherentist framework. It is compatible with a foundationalist conception of moral justification.

Here, in outline, is a more detailed account of how the method works. Again, we begin with intuitions about particular problems, particular cases. If our initial interest is in a problem about which we have no intuitions, or about which our intuitions are weak or conflicting, we should, as I have suggested with reference to the problem of abortion, find closely related cases about which we have confident intuitions and work from these. The question arises, however, why we should carry the inquiry any further. Why cannot we rest content with our intuitions, allowing ourselves to be guided by them on a case-by-case basis? Part of the answer, of course, is that there are many moral problems about which we have no intuitions, or about which the intuitions we have are weak, conflicting, or obviously suspect or dubious. We need a method for determining what we should believe and what we should do in cases such as these.

When one's moral intuition is challenged by another person, it is natural to respond by appealing to claims of a higher level of generality that imply or explain the intuition. The assumption is that the credibility of the intuition is enhanced if it can be subsumed under a plausible moral principle. So, for example, the intuitive judgment that it would be wrong to torture the cat for fun might be defended by

111

appealing to the principle that it is wrong to cause suffering without a justifying reason. Private moral reflection may follow the same dialectical pattern as moral disagreement between persons. We should challenge our own intuitions in much the way that an opponent might challenge them; but we may also respond in much the same way, by trying to bring them within the scope of a plausible principle.

But why suppose that the credibility of one's intuition is enhanced when it is shown that the intuition is implied by the conjunction of a moral principle and the facts of the case? One suggestion is that the principle may elucidate the intuition by identifying the features of the case that are morally salient. If, for example, one feels intuitively that it is wrong to kill animals for sport, one's objection is sharpened or focused if it is seen to follow from the more general view that it is morally objectionable, in the absence of sufficient justification, to deprive any individual of a good that that individual would otherwise have. The principle brings out more clearly at least part of what one finds intuitively objectionable.

No one supposes, of course, that just any principle will do. For the principle to support the intuition, it must have independent credibility. The principle itself may have intuitive appeal. According to coherentism, the *mere* fact that the principle implies the intuition provides some minimal epistemic support for each; for mutual coherence among beliefs is the criterion of justifiability. But, even according to coherentism, the principle will provide no more than token support for the intuition unless it is itself well integrated within a larger network of mutually coherent beliefs. Hence the method of reflective equilibrium demands that the principle itself be tested for consistency and coherence with other intuitions and principles. Its implications about particular cases should not conflict with one's intuitive judgments about those cases and, to the greatest extent possible, its implications should not contradict the implications of other principles one accepts. It is, of course, too much to require that moral principles not have conflicting implications: conflict is the price of pluralism. But conflicts should, in principle, be resoluble, in that one recognizes the necessity of one value's yielding to another, though not without some irreducible loss.

So the defense of one's initial intuition by subsuming it under a more general principle is only the beginning of moral inquiry. The principle must itself be assessed by testing its implications for consistency with one's other beliefs. One need not accept coherentism in order to appreciate the importance of this test. There are practical reasons why inconsistent moral beliefs are problematic: they may, for example, provoke indecision and, ultimately, paralysis of the will. More importantly, the achievement of greater coherence among one's beliefs diminishes the likelihood of error by helping one to identify and eliminate moral beliefs produced by self-interest, faulty reasoning, failure of imagination, illusory metaphysical beliefs, impaired faculties, and other sources of distorted or mistaken belief.

But there is a deeper basis for trying to subsume an intuition under a principle that is itself supported by its power to unify and explain a range of other intuitions. This is that the process of achieving increasing coherence among principles and intuitions facilitates the discovery of deeper values and also brings surface beliefs

112

about particular cases into alignment with those deeper values in a way that reveals and illuminates the connections between them. When one seeks to formulate a moral principle that implies and illuminates a widespread and robust intuition about a particular problem or case, one is in fact groping or probing for deeper values of which it is a surface manifestation. The expectation that the principle will illuminate and explain the intuition assumes that the intuition is in fact an expression, in a particular context, of a value that is deeper, more basic, and more general than the intuition itself. One's efforts to formulate the principle and to revise and refine it in a way that brings more and more common intuitions within its scope are attempts to capture or articulate some core principle in its full generality, to get its form exactly right, omitting nothing, however subtle.

This process, as I have described it, is indistinguishable from that endorsed by the coherentist practitioners of reflective equilibrium. One seeks support for an intuition by appealing to a principle, then one seeks to support the principle by demonstrating its compatibility with other intuitions, and so on. Why not understand the method as most people do, in coherentist terms?

There are various general objections to coherentist accounts of justification in ethics (e.g., Gaus 1996: ch. 6). I will not rehearse them here. (There are also, of course, general objections to foundationalism; I will not discuss those either.) I will simply note two problems that I find particularly disturbing. One is that, according to coherentism, no belief is immune to rejection, no matter how compelling it may be. If its elimination from the network of beliefs would enhance overall coherence within the network, the belief must go. Indeed, it seems possible, though not likely, that a coherentist approach to the pursuit of reflective equilibrium could lead ultimately to the rejection of every belief with which one started. Both these suppositions, however, are alien to moral life and moral reflection. There are some moral beliefs that we simply cannot give up just for the sake of greater coherence. Sometimes we must hold tenaciously to certain convictions even when it seems that greater coherence or systematicity could be achieved by rejecting it. According to the coherentist, one's moral beliefs are like pieces in a game: one shuffles them around, sacrifices some, and acquires others, all for the sake of achieving certain relations among them. No piece has any significance in itself; it has significance only in relation to the other pieces and in particular in the contribution that it makes to the whole of which it is a part. If moral reflection were really a game like this, in which our moral beliefs had no claim to commitment from us and thus were always expendable in the service of coherence, coherence would be easily achievable. It is because some of our moral beliefs compel our allegiance independently of their inferential relations to other beliefs that full coherence always seems such a distant goal.

A closely related worry about coherentism is that it assigns the same epistemic status to our intuitions about particular cases that it assigns to the deeper principles of which the intuitions are expressions. They stand in relations of reciprocal support: the principles imply the intuitions and we can therefore infer our way to the principles on the basis of the intuitions. But in fact the relations of reciprocal

support seem asymmetrical: the principles seem to be epistemically more basic, more secure. They articulate our core values which unify, explain, and justify our intuitive judgments. Our intuitions do not so much justify the principles as provide evidence of their existence and guidance as to their nature. In short, the principles are foundational with respect to the intuitions. Insofar as our intuitions are reliable sources of moral knowledge, they are so because they are expressions of, and point back to, a range of deeper, more general values.

As I noted earlier, foundationalism is distinguished by the view that some beliefs are justified independently of their relations to other beliefs. I will refer to these beliefs as "foundational." Among those who accept a foundationalist moral epistemology, there is a rough division between those who take certain nonmoral beliefs to be foundational and those who identify certain moral beliefs as foundational. Among the latter, there is a further division between those who take intuitions to be foundational (e.g., Ross 1930 and Gaus 1996) and those who take some general principle or principles to be foundational (e.g., Sidgwick 1907). In Geoffrey Sayre-McCord's words:

> Many have treated the privileged [that is, foundational] moral beliefs as roughly on a par with perceptual judgments and suggested that the justification of our various moral principles parallels the kind of justification our scientific principles receive from perception. Others have thought that our privileged moral beliefs concern, instead, the most general and abstract principles of morality, and that these in turn serve to justify (or not) our other beliefs deductively.
>
> (Sayre-McCord 1996: 150)

The view that I have suggested is of this second sort. But it is distinguished from many views of this sort in that it does not regard the foundational principles as self-evident or accessible directly through the exercise of intuition. Many philosophers (such as Sidgwick 1907 and Unger 1996) have regarded our intuitions about principles as more reliable than our intuitions about particular cases. But, insofar as a moral principle is substantive in character rather than merely formal (for example, "Treat like cases alike"), it seems a mistake to have confidence in our intuitive apprehension of the principle. To be justified in accepting a moral principle, we must first understand what it commits us to in particular cases. As William James noted in a letter written long before he became a practicing philosopher, "No one sees farther into a generalization than his own knowledge of the details extends" (Barzun 1983: 14). So, while I regard the principles rather than our intuitions as foundational, I do not think that moral inquiry can proceed by deducing conclusions about particular cases from self-evident moral principles. Rather, *the order of discovery is the reverse of the order of justification*. Although the deeper principles are explanatorily prior, we have to work our way to them via our intuitions in much the way that scientists work towards general principles via perceptual data. The process of discovering and formulating the more general principles is evidently difficult and intellectually demanding, rather in the way that

discovering the syntactic structures that govern our use of language is. As this familiar analogy suggests, as we grope our way towards the principles, we may be discovering what we antecedently believe, albeit below the level of conscious awareness, or discovering truths of which we were unaware. The principles that we hope to uncover may express deep dispositions of thought and feeling that operate below the level of consciousness to regulate our intuitive responses to particular cases.

If this is right, it explains why we experience the process of moral inquiry as a process of discovery rather than an exercise of choice or will. It also explains why the foundationalist approach I am describing should coincide with coherentism in holding that we may expect to arrive at a moral theory, if at all, only near the end of reflection about particular problems and cases rather than coming to the problems with a theory ready to be "applied." It explains why it is suspect when philosophers emerge from graduate school already believing themselves to be in possession of the correct moral theory.

Challenges

This brief sketch of an account of moral justification raises far more questions than can be answered, or even addressed, in the remainder of this short essay. The central question is, of course, what reason is there to suppose that moral intuitions, even those that are widespread and robust, are reliable guides to the formulation or discovery of principles that in turn provide reliable foundations for particular moral judgments? If the principles toward which common intuitions seem to direct us turned out to be luminously self-evident, this could reinforce our confidence in the intuitions. But that seems unlikely. Thus far the efforts of philosophers to find general principles that unify and provide a deeper foundation for common intuitions have led them to principles that have little immediate appeal on their own – for example, the principle that the intention with which one acts, or the causal relations among the consequences of one's act, matters to the permissibility of the act. Unlike the principle that one ought to do the act that would have the best consequences, impartially considered (which may seem plausible *until* one tests its implications against common intuitions), these principles seem attractive primarily because they promise to provide some underlying unity to our intuitions. They may not seem, therefore, to provide independent support for the intuitions, apart from showing how they cohere with one another.

There are, moreover, numerous objections to according any epistemic authority to moral intuitions, however common or robust. Perhaps the easiest of these objections to rebut is that there is an "inbuilt conservatism" (Singer 2005: 347) in any method of moral justification that gives an important role to moral intuitions. This concern goes back at least to Mill, who, in Alan Ryan's words, thought that reliance on intuitions in ethics "amounted to the sanctification of any opinion

that had been held long enough and deeply enough. It was the great intellectual buttress of social, moral, political, and intellectual conservatism" (Ryan 2011: 62). Yet the fact that each of us has many intuitions that, on inspection, conflict with others we hold with equal conviction, so that even common and robust intuitions frequently conflict with other equally common and robust intuitions, means that our efforts to achieve consistency among common intuitions are inevitably revisionist, sometimes radically so, when strong intuitions cannot be reconciled. Philosophers who accept that moral intuitions have a presumptive epistemic authority, and thus have employed the method I described in the earlier section "A Sketch of a Foundationalist Conception of Moral Justification," have consequently been driven by the demand for consistency to defend quite heterodox and indeed counterintuitive conclusions, such as that infanticide can be permissible in a variety of circumstances, that each person in wealthy societies ought to give the bulk of his or her wealth to people in impoverished societies, and so on.

Another familiar objection to the appeal to moral intuitions is that they are often tainted by their origin. Earlier I quoted Singer as observing that many common intuitions have their ultimate source in primitive religious beliefs, ancient taboos about sex, and social practices that were useful in a world that is no longer ours. There are many other tainted sources, such as superstitions concerning purity and defilement, and, perhaps most important, individual and collective self-interest. The vast majority of whites in the antebellum South thought it obvious that the enslavement of blacks was morally justified. Although they sought biblical and biological warrant for the practice, what really motivated their belief was crude self-interest. As long as they all had a strong interest in maintaining the practice and could reinforce each other's beliefs by participating in the practice and raising their children to accept it as part of the natural background to their lives, they were able to insulate their intuitive sense of the rectitude of the institution of slavery from challenges that would otherwise have disturbed it. People do the same today with their belief that it is permissible to kill animals in order to eat them. Because eating meat gives them pleasure, people assume that it is in their interest (though in the forms and quantities in which they consume it, it is not). Most people therefore eat meat and this itself shields them from critical reflection. For they assume that because virtually everyone does it, including the very nicest people they know, it simply cannot be seriously wrong to do it. Yet without the blinkering effects of self-interest, the many powerful moral objections to this practice would be obvious.

The fact that many common moral intuitions are discredited by their provenance is not, however, a problem if we have rejected the idea that intuitions result from the exercise of some special faculty for the direct inspection of moral reality. No one supposes that all moral intuitions, or even all "considered moral judgments," are correct. They are instead merely *appearances* (Huemer 2005: 99–105), some of which, we recognize, are bound to be delusions. According to the method of moral inquiry I described earlier, an essential part of the process of arriving at

foundational moral principles is the filtering out of intuitions that are contaminated at the source.

Some have argued, however, that the eradication of bias in our intuitions is impossible because recent work on the psychology of moral judgment and in "experimental philosophy" has revealed sources of systematic bias and irrationality in our moral intuitions that seem inherent in our psychological nature. This work has shown, for example, that our retributive intuitions become stronger and harsher in the presence of a bad smell, which is obviously irrelevant to whether or to what extent a person deserves to be harmed. Given that our intuitions are subject to this kind of distortion, it seems that they cannot be reliable guides to the discovery of foundational moral principles. For one aim of the method of moral justification I have described is to abstract from or transcend the idiosyncrasies of our personal psychologies or subjective points of view in order to achieve the greatest possible degree of objectivity and impartiality. Yet if common intuitions are systematically distorted by psychological mechanisms of which we are unaware, and that may not even be accessible to introspection, then we have presumably been failing to achieve objectivity even while doing all we can to succeed.

What this new work in psychology and experimental philosophy shows, however, is that we have not been doing all we could. For until just recently we have not been doing the invaluable work that psychologists and experimental philosophers have begun to do in identifying these subtle sources of distortion and irrationality in common moral intuitions. Even if these distorting influences cannot be eradicated in our intuitive thinking, they can, once they have been recognized, be controlled for in our moral theorizing. What experimental philosophers have forced us to see is that their empirical investigations are essential to the process of filtering raw intuitions and hence indispensable in the process of moral justification I have described.

A third concern about the epistemic status of moral intuitions arises, paradoxically, from their surprising uniformity over time and across cultures. We have been impressed for so long by the claims of cultural anthropologists, postmodern relativists, undergraduates, and others about the diversity of moral opinion that we have tended to overlook how much agreement there really is. Interestingly, moral disagreements seem to widen and intensify the more we abstract from particular cases and focus instead on matters of principle or theory. When the partisans of different schools of moral thought turn their attention to particular cases, there is far more intuitive agreement than their higher-level disputes would lead one to expect. As teachers of ethics will testify, there is often a remarkably high level of agreement among students about what it is permissible to do in specially contrived examples, such as the well-known trolley cases (which have figured heavily in much of the recent work on the psychology of moral judgment), even though the students may come from widely varying religious and cultural backgrounds and have never had any experience of problems of the sort about which they are invited to offer an intuitive judgment. What accounts for the agreement?

A potentially debunking explanation is that it has a biological basis. It may be, as Singer puts it, that "our common evolutionary heritage has, unsurprisingly, given us a common set of intuitive ideas about right and wrong" (Singer 2005: 349). Some theorists have argued that human beings have evolved a special faculty of moral judgment that functions in a way parallel to the way innate syntactical structures function in governing our use of language. Colin McGinn attributes this conception of moral knowledge to Noam Chomsky, who originally formulated the parallel account of our knowledge of language. McGinn writes:

> According to Chomsky, it is plausible to see our ethical faculty as analogous to our language faculty: we acquire ethical knowledge with very little explicit instruction, without great intellectual labour, and the end-result is remarkably uniform given the variety of ethical input we receive. The environment serves merely to trigger and specialise an innate schematism. Thus the ethical systems of different cultures or epochs are plausibly seen as analogous to the different languages people speak – an underlying universal structure gets differentiated into specific cultural products.
>
> (1993: 30)

Common moral intuitions are the deliverances of this biologically innate moral faculty. (For a recent, elaborately detailed working out of this idea, see Mikhail 2011.)

One concern about this hypothesis, however, is that there seems to be no reason to suppose that natural selection would have produced in us a faculty of judgment that would track moral truth rather than reproductive advantage. Another possibility, of course, is that the psychological capacities we have evolved for other purposes enable us to reason not only about matters relevant to survival and reproduction but also about moral matters, and the substantial uniformity of moral intuition over time and across cultures is the result of our applying the same capacities to the same object – namely, moral reality.

Some philosophers, however, have argued that it is explanatorily superfluous to suppose that our intuitions are responses to an independent moral reality when evolutionary theory has already supplied a plausible account of their origin in the mechanisms of natural selection. It seems uncontroversial, for example, that common intuitions about parental responsibility, marital fidelity, loyalty, promise-keeping, and so on have an evolutionary basis. Evolutionary biologists and psychologists have produced elegant explanations of various forms of moral behavior by reference to such notions as inclusive fitness and reciprocal altruism, and one can be confident that more and more instances of moral behavior and belief will come within the scope of biological explanation as our understanding advances. It may therefore seem more plausible to regard our intuitions as the products of natural selection than as glimpses of moral reality.

Yet many people have certain intuitions that are contrary to those that evolutionary theory would predict they would have. It is difficult, for example, to find any reproductive advantage in the conviction of early abolitionists that slavery was

wrong, or in the intuition that it is in general wrong to harm or kill nonhuman animals, or in the belief that we have exacting duties to aid impoverished people in other countries, to reduce our own fertility to control population growth, and to make other sacrifices in our own quality of life to avoid adversely affecting that of future people. Generally such beliefs are also, at least initially, at variance not only with individual self-interest but also with the received beliefs, including the religious beliefs, within the culture of those who have them. Yet these beliefs often spread, though slowly and more by force of example than through persuasive reasoning. They are, moreover, often genuine intuitions rather than the products of inferential reasoning – as, for example, in the case of a child who, intuitively repulsed at the discovery of what meat really is and how it is produced is prompted to demand to be allowed to become a vegetarian.

That people have such beliefs, and that in time they often come to be widely regarded as evidence of moral progress, supports a conception of moral intuition that is compatible with moral realism; namely, that some intuitions – those that survive both the initial process of filtering and the testing for consistency – are judgments that are true and that direct us to the discovery of foundational moral principles that are also true.

Acknowledgments

References

Barzun, Jacques (1983) *A Stroll with William James*, New York: Harper & Row.

Brandt, Richard B. (1979) *A Theory of the Right and the Good*, Oxford: Clarendon Press.

Gaus, Gerald (1996) *Justificatory Liberalism: An Essay on Epistemology and Political Theory*, New York: Oxford University Press.

Hare, R.M. (1975) "Abortion and the Golden Rule," *Philosophy and Public Affairs* 4: 201–22.

Huemer, Michael (2005) *Ethical Intuitionism*, Basingstoke, UK: Palgrave Macmillan.

McGinn, Colin (1993) "In and Out of the Mind," *London Review of Books* (2 December): 30–1.

Mikhail, John (2011) *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*, Cambridge: Cambridge University Press.

119

Rachels, James (1986) *The End of Life: Euthanasia and Morality*, Oxford: Oxford University Press.

Rawls, John (1972) *A Theory of Justice*, Oxford: Clarendon Press.

Ross, W.D. (1930) *The Right and the Good*, Oxford: Clarendon Press.

Ryan, Alan (2011) "The Passionate Hero, Then and Now," *New York Review of Books* (December 8): 59–63.

Sayre-McCord, Geoffrey (1996) "Coherentist Epistemology and Moral Theory," in *Moral Knowledge?* eds. Walter Sinnott-Armstrong and Mark Timmons, New York: Oxford University Press, pp. 137–89.

Sidgwick, Henry (1907) *The Methods of Ethics*, 7th edn, London: Macmillan.

Singer, Peter (1974) "Sidgwick and Reflective Equilibrium," *The Monist* 58: 490–517.

Singer, Peter (1993) *Practical Ethics*, 2nd edn, Cambridge: Cambridge University Press.

Singer, Peter (2005) "Ethics and Intuitions," *Journal of Ethics* 9: 331–52.

Timmons, Mark (1987) "Foundationalism and the Structure of Ethical Justification," *Ethics* 97: 595–609.

Unger, Peter (1996) *Living High and Letting Die: Our Illusion of Innocence*, New York: Oxford University Press.

Further Reading

Berker, Selim (2009) "The Normative Insignificance of Neuroscience," *Philosophy and Public Affairs* 37: 293–329.

Daniels, Norman (1996) *Justice and Justification: Reflective Equilibrium in Theory and Practice*, Cambridge: Cambridge University Press.

DePaul, Michael R. (1993) *Balance and Refinement: Beyond Coherence Methods of Moral Inquiry*, London: Routledge.

Greene, Joshua D. (2008) "The Secret Joke of Kant's Soul," in *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotional, Brain Disorders, and Development*, ed. Walter Sinnott-Armstrong, Cambridge, MA: MIT Press, pp. 35–79.

Lillehammer, Hallvard (2011) "The Epistemology of Ethical Intuitions," *Philosophy* 86: 175–200.

Street, Sharon (2006) "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127: 109–66.