

15

Moral Intuitionism Meets Empirical Psychology

Walter Sinnott-Armstrong

... if this be all, where is his ethics? The position he is maintaining is merely a psychological one.

(Moore 1903: §11, 11)¹

G. E. Moore's diatribe against the naturalistic fallacy in 1903 set the stage for most of twentieth-century moral philosophy. The main protagonists over the next sixty years were intuitionists and emotivists, both of whom were convinced by Moore that empirical science is irrelevant to moral philosophy and common moral beliefs. Even in the 1970s and 1980s, when a wider array of moral theories entered the scene and applied ethics became popular, few moral philosophers paid much attention to developments in biology and psychology.

This isolation must end. Moral philosophers cannot continue to ignore developments in psychology, brain science, and biology. Of course, philosophers need to be careful when they draw lessons from empirical research. As Moore and his followers argued, we should not jump straight from descriptive premises in psychology or biology to positive moral conclusions or normative conclusions in moral epistemology. That would be a fallacy.² Nonetheless, psychology can still affect moral philosophy in indirect ways. That is what I want to illustrate here. I will trace an indirect path from empirical premises to a normative conclusion

For comments on drafts and oral presentations, I thank Robert Audi, Paul Bloomfield, George Pappas, Bernard Gert, Terry Horgan, Don Loeb, James Moor, Matthew Nudds, Diana Raffman, John Skorupski, Roy Sorensen, and others in audiences at the Ohio State University, the University of Connecticut, the University of Vermont, the University of Nebraska at Lincoln, and the University of Edinburgh.

¹ Here Moore was discussing the claim that the object of desire is not pleasure, but his charge against naturalism extended much further.

² Although some such arguments are formally valid. See my 2000: 159–74. I add the qualification 'positive' because 'Bertie and Madeleine are dead' might entail 'It is not the case that Bertie ought to marry Madeleine.'

in moral epistemology. In particular, I will argue that some recent research in psychology and brain science undermines moral intuitionism.

1. What is moral intuitionism?

Some philosophers define moral intuitionism as the structural view that there are many moral values or requirements with no systematic unification or ranking. Other philosophers see moral intuitionism as the metaphysical view that moral properties are non-natural. Neither of these views concerns me here. I mention them only to set them aside.

The kind of moral intuitionism that is my target here is a position in moral epistemology, which is general epistemology applied to moral beliefs. The deepest challenge in moral epistemology, as in general epistemology, is raised by a skeptical regress argument: Someone is justified in believing something only if the believer has a reason that is expressible in an inference with premises that the believer is already justified in believing. This requires a chain of inferences that must continue infinitely, close into a circle, or stop arbitrarily. Academic skeptics reject all three options and conclude that there is no way for anyone to be justified in believing anything. The same regress arises for moral beliefs (cf. Sinnott-Armstrong 1996: 9–14; 2002*a*).

The simplest way to stop this regress is simply to stop. If a believer can work back to a premise that the believer is justified in believing without being able to infer that premise from anything else, then there is no new premise to justify, so the regress goes no further. That is how foundationalists stop the regress in general epistemology. Moral intuitionists apply foundationalism to moral beliefs as a way to stop the skeptical regress regarding moral beliefs.

The motivation behind moral intuitionism is not always to stop the skeptical regress,³ but that use of moral intuitionism is common and is what concerns us here, so we can use it to decide among possible definitions of moral intuitionism. What we need to define is the weakest version of moral intuitionism

³ As George Pappas reminded me, part of foundationalism can be separated from the denial of skepticism. Foundationalists can be skeptics if foundationalism claims only that a believer is justified in a belief only if the belief is either non-inferentially justified or inferrable from non-inferentially justified beliefs. Non-skeptical foundationalists merely add that some beliefs are justified. Analogously, moral intuitionism could be seen as a claim about the structure of justification separate from any denial of skepticism. Nonetheless, I will define moral intuitionism to include the denial of moral skepticism because almost all actual moral intuitionists do deny skepticism and because I am concerned with whether moral intuitionism can succeed as a response to skepticism.

that is strong enough to solve the regress problem that would lead to moral skepticism. Here it is:

Moral intuitionism is the claim that some people are adequately epistemically justified in holding some moral beliefs independently of whether those people are able to infer those moral beliefs from any other beliefs.⁴

Several features of this definition are worth highlighting.

So defined, moral intuitionism is not about knowledge. It is about justified belief. This makes it normative. Psychologists sometimes define intuitionism as a descriptive claim about the nature and origins of moral beliefs (see Haidt 2001). Such descriptive claims are not intended to stop the skeptical regress; so they do not concern me here.

More specifically, the defining claim of moral intuitionism is about what is *epistemically* justified because moral skeptics win if the only justification for holding moral beliefs is that those belief states have beneficial practical effects. Similarly, moral skeptics win if some moral beliefs are inadequately justified but none is *adequately* justified, that is, justified strongly enough that the believer ought to believe it as opposed to denying it or suspending belief.⁵ Accordingly, I will henceforth use 'justified' as shorthand for 'adequately epistemically justified'.

To say that moral believers are justified independently of an inferential ability is just to say that they would be justified even if they lacked that ability, that is, even if they were not able to infer those beliefs from any other beliefs. This independence claim can hold even when moral believers are able to infer those moral beliefs from other beliefs as long as they do not need that inferential ability to be justified.⁶ This notion of need will become prominent later.

⁴ Although I define moral intuitionism in terms of 'some' believers and 'some' moral beliefs, all actual moral intuitionists claim that a significant group of believers and beliefs can be justified non-inferentially. It also might seem odd that a theory counts as moral intuitionism on my definition if it holds that some beliefs based on testimony are justified independently of any inference or inferential ability. However, my arguments will apply to such views, so I see no pressing need to complicate my definition so as to avoid these problems.

⁵ I will discuss pro tanto justifiedness in responses to objections below. For more detail on kinds of justifiedness, see my 2002*b*: 17–25.

⁶ Compare Ross 1930: 29: 'without any need of proof'. (Since not all inferences are proofs, and Ross does not mention abilities, he might not deny the need for an inferential ability.) Contrast Moore (1903: 77), who sees moral intuitions as unprovable. Moore's stronger claim is not needed to stop the skeptical regress. Notice also that opponents of moral intuitionism, who claim that an inferential ability is *necessary*, do not have to claim that any inferential ability is *sufficient* to make any moral belief justified. At least the moral belief must also be *based on* the inferential ability. Other necessary conditions might also have to be met. Opponents of moral intuitionism can hold that an inferential ability is needed without specifying what, if anything, else is needed and, so, without specifying what is sufficient for justified moral belief.

I infer a belief when I go through a reasoning process of which the belief is the (or a) conclusion and other beliefs are premises. A believer is able to draw such an inference when the believer has enough information to go through a reasoning process that results in this belief if he had enough incentive and time to do so. This ability does not require self-consciousness or reflection about the beliefs or abilities. All that is needed, other than general intelligence, is for the requisite information to be encoded appropriately in the believer's brain at the time of belief.⁷

Some moral intuitionists claim only that certain moral beliefs are justified independently of any actual inference. However, that weak moral intuitionism is not enough to stop the skeptical regress. Even if whether certain moral beliefs are justified does not depend on any actual inference, it still might depend on the believer's ability to infer them from other beliefs. The ability to draw an inference cannot make a belief justified if beliefs in the inference's premises are not themselves justified. This requirement is enough to restart a skeptical regress. Thus, to meet the skeptical challenge, moral intuitionists must make the strong claim that some moral believers are adequately epistemically justified in holding some moral beliefs independently of any ability to infer the moral belief from any other belief.⁸ So that's what they claim.

Although this claim is strong, it has many defenders. Rational intuitionists see basic moral beliefs as analogous to beliefs in mathematical axioms, which are taken to be justified independently of inference. Moral sense theorists assimilate particular moral beliefs to perceptual beliefs, which are supposed to be justified independently of inference. More recently, reliabilists hold that any belief is justified if it results from a reliable process, regardless of whether that process has anything to do with any inference. I group these views together under my broad definition of moral intuitionism because my arguments will apply to them all.

Under my definition, there are at least two ways to deny moral intuitionism. Moral intuitionists claim that some moral believers would be justified even if they did not have any ability to infer their moral beliefs from any other beliefs. Some opponents object that moral beliefs always depend on some inference or inferential ability. However, the evidence (cf. Haidt 2001) strongly suggests

⁷ The relevant notion of ability, then, is not the same as when I am able to become justified in believing that there are ten coins in my pocket because I could take them out of my pocket and count them. To have an ability of the relevant kind, I must be able to infer the belief from other beliefs that I already have without gaining any new information. I hope that it is also clear that, when I write about needing 'an inferential ability', I am not referring to a general ability to draw just any old inference, or any inference of a certain form. What is at issue is the ability to infer the specific moral belief from other beliefs.

⁸ The stronger claim is also needed for moral intuitionism to contrast with its traditional opponent, moral coherentism, since coherentists do not claim that believers must actually draw inferences in order to be justified.

that people often have moral beliefs that do not result from any actual inference. It is harder to tell whether any moral beliefs are independent of any ability to infer. Nonetheless, I will grant for the sake of argument that some moral beliefs are spontaneous in the sense that they are independent of any inference or inferential ability.

Other opponents of moral intuitionism deny that moral believers are ever justified in holding such spontaneous moral beliefs if they lack certain inferential abilities. This conclusion follows if inferential abilities are always needed for a moral believer to be justified. That is what I will try to show.

2. When is confirmation needed?

We cannot answer this question directly. If a moral intuitionist baldly asserts that we do not need inferential abilities to back up our spontaneous moral beliefs, then this assertion begs the question. Similarly, if a critic of moral intuitionism baldly asserts that we do need inferential abilities to back up our spontaneous moral beliefs, then this assertion also begs the question. Neither side can win so easily. We need a less direct method.

One alternative uses analogies to non-moral beliefs. This path is fraught with peril, but it might be the only way to go. What this approach does is appeal to non-moral cases to develop principles of epistemic need and then later apply those principles back to moral beliefs. Let's try it.

I will formulate my principles in terms of when confirmation is needed, but they do not claim that the believer needs to go through any process of confirming the belief. The point, instead, is only that some confirmation needs to be available at least implicitly as information stored somehow in the believer that gives the believer an ability to infer the belief from some other beliefs.⁹ The question is when some such confirmation is needed in non-moral cases.

Suppose that I listen to my daughter in a piano competition. I judge that she played great and her rival was mediocre. Am I justified in trusting my judgment? Not if all I can say is, 'Her performance sounded better to me.' I am too biased for such immediate reactions alone to count as evidence. I still might be justified, if I am able to specify laudable features of her performance, or if I know that others agree, but some confirmation seems needed. Generalizing,

Principle 1: confirmation is needed for a believer to be justified when the believer is partial.

⁹ Confirmation need not always be evidence, since I want confirmation to include defeater defeaters, that is, reasons to discount what would otherwise keep a belief from being justified.

This principle also applies to direct perceptual judgments, such as when I believe that my daughter played middle C at just the right time in the midst of her piece. This partly explains why we prefer umpires, referees, and judges not to be parents of competitors. Even reliabilists can admit this principle because partiality often creates unreliability.

Second, imagine that each of us adds a column of figures, and I get one sum, but you get a different sum. Maybe I would be justified in believing that I am right if you were my child and I was helping you with your homework. However, if you are just as good at arithmetic as I am, then, when we get different answers, we need to check again to find out who made a mistake before either of us can be justified in believing that his or her answer is the correct one. We owe each other that much epistemic respect. The best explanation of this natural reaction seems to be

Principle 2: confirmation is needed for a believer to be justified when people disagree with no independent reason to prefer one belief or believer over the other.

This principle also applies when the person on the sidewalk looks like Tom Cruise to me but not to you. If I have no reason to believe that I am better than you at this identification, then I am not justified in believing that your belief is incorrect or that mine is correct.

A third principle concerns emotions. When people get very angry, for example, they tend to overlook relevant facts. They often do not notice excuses or apologies by the person who made them angry. We should not generalize to all emotions, but we can still endorse something like this:

Principle 3: confirmation is needed for a believer to be justified when the believer is emotional in a way that clouds judgment.

This explains why jurors are dismissed from a case that would make them too emotional. This principle applies even if their emotions do not bias them towards either side, so it is distinct from Principle 1, regarding partiality.

Next consider illusions. At least three kinds are relevant here. First, some illusions are due to context. Objects look larger when they are next to smaller objects, and they look smaller when they are next to larger objects. Since our estimates of their sizes are affected by their surroundings, we are not justified in trusting our estimates until we check their sizes in other circumstances or by other methods.

A second group of illusions arises from generalizations. For example, an oval that is shaded on top looks concave, but an oval that is shaded on the bottom looks convex. The explanation seems to be that our cognitive apparatus

evolved in circumstances where the light usually came from above, which would produce a shadow on the top of a concave oval (such as a cave opening) and on the bottom of a convex oval (such as an egg). Since we often overextend generalizations like this, we are not justified in trusting beliefs that depend on such generalizations until we check to determine whether our circumstances are exceptional.

The third kind of illusion involves heuristics, which are quick and simple decision procedures. In a passage with a thousand words, how many seven-letter words have the form '____in_'? How many seven-letter words have the form '____ing'? Most people estimate more words have the latter form, although that is impossible, since every word of the form '____ing' also has the form '____in_'. Why do people make this simple mistake? They seem to test how likely something is by trying to imagine examples and guessing high if they easily think of lots of examples. This is called the availability heuristic (Kahneman *et al.* 1982). Most people easily produce words that end in 'ing', but they have more trouble coming up with words that end in 'in_' because they do not think of putting 'g' in the last place. In cases like this, the availability heuristic is misleading. Accordingly, they do not seem adequately epistemically justified in trusting beliefs based on such heuristics until they check on whether they are in circumstances where the heuristics work.

This quick survey of three common kinds of illusion suggests

Principle 4: confirmation is needed for a believer to be justified when the circumstances are conducive to illusion.

This principle would apply as well to many other kinds of illusions.

A fifth and final principle considers the source of a belief. If you believe that George Washington never told a lie, and if this belief comes from a legend spread by Washington's allies to gain power, then you are not justified in believing the legend, though it still might be true. Even if you believe only that Washington was unusually honest, this belief might be a lingering effect of this childhood story, and then its origin makes this belief need confirmation. The point can be generalized into something like

Principle 5: confirmation is needed for a believer to be justified when the belief arises from an unreliable or disreputable source.

This principle also explains why we do not view people as justified in beliefs based only on prejudice and stereotypes.

These five principles, although distinct, complement each other. When a belief is partial, controversial, emotional, subject to illusion, and explicable by dubious sources, then all of these principles apply. In such cases, they work

together to make it even clearer that confirmation is needed for justified belief. Even if not all of these principles apply, the more that do apply, the clearer it will be that there is more need for more confirmation. We might think of this as a sixth principle.

I do not claim that these principles are precise or that my list is complete.¹⁰ What I do claim is that these principles or some close relatives seem plausible to most people and are assumed in our shared epistemic practices. I also claim that they make sense because they pick out features that are correlated with reliability and other epistemic values.

Most importantly, I claim that these principles apply in all areas of belief. My illustrations include beliefs about arithmetic, language, history, identity, value, sound, size, and shape, but the same principles apply in scientific research, religion, and so on. The main question here is whether they apply to moral beliefs. Admittedly, morality might be a special case where these principles do not apply. However, unless someone can point to a relevant difference between these other areas and moral beliefs, it seems only fair to apply these same standards to moral beliefs when asking whether moral beliefs are justified. So that's what I will do.

3. When are moral beliefs justified?

Some of these principles can be applied only with the help of empirical research. Others are easier to apply. Let's start with the easy ones.

Partiality

Principle 1 says that partiality adds a need for confirmation. But what is partiality? A judge is often called partial when the judge's self-interest is affected by the outcome of the case. However, if the judge's self-interest does not influence the judge's decision, so the judge would have made the same decision if the judge's self-interest had not been involved, then it is natural to say that the judge's decision is not partial, even if the judge is partial. Analogously, believers can be called partial whenever their beliefs affect their self-interest either directly or indirectly (by affecting the interests of people whom they care about). Beliefs are then partial only when the believer's self-interest influences whether the believer holds that belief. Thus, a partial believer can hold an

¹⁰ One additional principle might claim that confirmation is needed when errors are costly. This principle applies to moral beliefs insofar as moral errors are costly.

impartial belief (or can hold it impartially) if the believer has an interest in holding the belief, but that interest does not influence whether the believer holds that belief.

Partiality of a belief can't be all that triggers Principle 1. To see why, recall the examples that motivated Principle 1. Because I am biased in favor of my daughter, even after watching her play in the piano competition, I need confirmation to be justified in believing that my daughter played better than her rival. Maybe my interest in her victory did not influence my assessment, but the danger of such influence is enough to create a need for confirmation. Admittedly, if I can rule out such influence, then I can be justified in believing that my daughter played better, but the only way to rule out such influence involves independent confirmation. Thus, confirmation seems needed when the believer is partial, even if the believer is not actually influenced by that partiality, so the belief is not partial. Since confirmation is also needed when the belief is partial, Principle 1 requires confirmation when either the believer or the belief is partial.

To apply this principle to moral beliefs, we need to determine whether moral beliefs affect our self-interest either directly or indirectly. The answer seems clear: Moral beliefs affect us all. It can be very expensive to believe that we are morally required to help the needy, but it can be even more expensive if others do not believe that they are morally required to help us when we are in need. It can also cost a lot to believe that we have to tell the truth or to keep a promise. And all of us know or should know that, if killing, stealing, lying, cheating, and promise-breaking were generally seen as morally permitted, then we would be more likely to get hurt by others doing such acts. Life would be more 'solitary, poor, nasty, brutish, and short', as Hobbes put it. Moreover, if an individual did not see such acts as immoral, then he or she would be more likely to do them and then to be punished in various ways—'if not by law, by the opinion of his fellow creatures', as Mill said. Special interests also arise in special cases: Women and men know or should know that, if abortion is not seen as morally permissible, then they, their friends, or their daughters will be more likely to suffer more. Moral beliefs about affirmative action affect the interests of the preferred groups and also the non-preferred groups. And so on. Indeed, on many views, what makes an issue moral in nature is that interests are significantly affected by the judged actions. Since moral beliefs about actions affect those actions, our moral beliefs themselves affect our interests at least indirectly. Finally, social groups often form around and then solidify moral beliefs (cf. Chen and Tetlock *et al.* as discussed in Haidt 2001). People who believe that homosexuality is immoral find it harder to get along with homosexuals and easier to get along with homophobes. Conversely, people who

believe that homosexuality is not immoral find it harder to get along with homophobes and easier to get along with homosexuals. Some might try to fake moral beliefs in order to get along, but few of us are good enough actors, and those who believe that homosexuality is immoral usually also believe that they ought not to pretend otherwise in order to get along with homosexuals. Thus, our moral beliefs affect our social options as well as our actions.

Many moral beliefs might seem to have no effect on us. If I believe that it was immoral for Brutus to stab Caesar, this moral belief by itself will not change my social options or rule out any acts that I could do today. Still, given universalizability, my judgment of Brutus seems to depend on a principle that does apply in other cases where my self-interest is involved more directly. Arguments from analogy also might force me to take a moral stand that affects my interests. Thus, any moral belief can affect my self-interest indirectly.

Because our moral beliefs affect our self-interest so often in so many ways at least indirectly, we cannot be justified in assuming that any of us is ever fully impartial as a moral believer. Even if our self-interest is not involved in some exceptional case, we still need a reason to believe that our self-interest is not affected in that case, since we know or should know that such effects are very common and often hidden. The facts that partiality is so common in this area and so difficult to detect in ourselves are what create a need for confirmation of all moral beliefs, according to Principle 1.

Disagreement

Principle 2 says that disagreement creates a need for confirmation. Many people seem to think that this principle is easy to apply to moral beliefs because moral disagreement is pervasive. In their view, people from different cultures, time periods, social classes, and genders disagree about a wide variety of particular moral judgments and general moral principles.

Actually, the extent of moral disagreement is not obvious. One reason is that people who seem to disagree are often judging different actions or using different concepts. Also, many apparently moral disagreements are really factual, since those who seem to disagree morally would agree in their moral judgments if they agreed about the facts.

Still, straightening out concepts and non-moral facts seems unlikely to resolve all apparently moral disagreements. One reason is that people often express different moral beliefs about hypothetical cases where all of the facts are stipulated, so these moral believers seem to accept the same non-moral facts. Admittedly, descriptions of these situations usually leave out important facts, and moral believers might interpret the hypothetical cases in light of different

background beliefs. But there still seem to be lots of cases where all relevant non-moral facts are agreed upon without leading to agreement in moral belief.

This claim is supported by a study in which Jana Schach Borg and I surveyed fifty-two undergraduates at Dartmouth College, using thirty-six scenarios, including the well-known side-track and fat-man trolley cases. In both cases, five people tied to a track will be killed by a runaway trolley if you do nothing, and the only way to save the five is to kill one other person.¹¹ In the side-track version, you can save the five only by pulling a lever to divert the trolley onto a side-track where it will run over one victim. In the fat-man variation, you can save the five only by pushing a fat man in front of the trolley so that his body will stop the trolley before it hits the five. In two rounds, 35 per cent then 43 per cent of our subjects said that it would be wrong to divert the trolley onto the side-track. When the same scenario was described with more vivid language, 61 per cent then 45 per cent judged diversion wrong in the two rounds. In contrast, 76 per cent in the first round then 88 per cent in the second round judged it wrong to push the fat man. (Interestingly, there were still 35 per cent in the first round and 18 per cent in the second round who said that they would push the fat man.) These percentages did not change much with more vivid language. Thus, we found significant disagreement about the very cases that philosophers often cite to support their theories.¹²

There is, admittedly, more agreement about other cases: Would it be wrong to push the fat man in front of the trolley just because you are angry at him for beating you at golf when killing him will not save or help anyone else? I hope and expect that 100 per cent would answer, 'Yes.' But what would that show? The universality of moral beliefs about cases like this one could hardly be used to justify any moral theory or any controversial moral belief.¹³ Such cases cannot get moral intuitionists all that they seem to want.¹⁴

Moral intuitionists might respond that all they claim is that *some* moral beliefs are non-inferentially justified. One case seems enough to establish that

¹¹ These cases originate from Foot 1967. If such cases seem unrealistic, see the real case at www.cnn.com/2003/US/West/06/21/train.derail.ap/index.html.

¹² For more evidence of disagreement, see Haidt *et al.* 1993.

¹³ Compare Descartes's 'I think,' which is nowhere near enough to ground science. Notice also that I am talking about actual moral beliefs, not possible moral beliefs. There might be an infinite number of possible moral beliefs that would garner agreement from everyone who understands them. However, there still might be a high rate of disagreement among the actual moral beliefs that people bother to form. That is what matters when we ask whether our actual moral beliefs are justified.

¹⁴ Some intuitionists might claim agreement on qualified general principles, such as that it is morally wrong to kill anyone in a protected class without an adequate reason. Of course, those who accept this formula still disagree about which class is protected and which reasons are adequate. Similarly, although many people (today!) agree that all moral agents deserve respect, different people count different acts as violating that rule by showing disrespect. It is not clear whether to count agreement on such indeterminate formulas as real moral agreement. See Snare 1980.

claim. However, the fact that there is so much disagreement in other cases affects the epistemology of cases where there is no disagreement. Compare a box with one hundred thermometers. We know that many of them don't work, but we are not sure how many. If we pick one thermometer arbitrarily from the box, and it reads 77 degrees, then we are not justified in believing that the temperature really is 77 degrees, even if we were in fact lucky enough to pick a thermometer that works. Of course, if we confirm that this thermometer works, such as by testing it against other thermometers, then we can use it to form justified beliefs, but we cannot be justified in trusting it before we confirm that it works.¹⁵ Similarly, if we know that many moral intuitions are unreliable because others hold conflicting intuitions, then we are not justified in trusting a particular moral intuition without some reason to believe that it is one of the reliable ones. If we know that everyone agrees with that particular moral intuition, then we might have reason to trust it. But that is just because the known agreement provides confirmation, so it does not undermine the point that some confirmation is needed, as Principle 2 claims.

Emotion

Next consider Principle 3, which says that emotions that cloud judgment create a need for confirmation. It is hard to tell whether this principle applies to moral beliefs. Philosophers and others have argued for millennia about whether moral beliefs are based on emotion or on reason. They also argue about which emotions, if any, cloud judgment. How can we resolve these debates? Luckily, some recent empirical studies suggest an answer.

Haidt and his group have been accumulating an impressive body of behavioral evidence for what they call the social intuitionist model:

This model suggests that moral judgment is much like aesthetic judgment: we see an action or hear a story and we have an instant feeling of approval or disapproval. These feelings are best thought of as affect-laden intuitions, as they appear suddenly and effortlessly in consciousness, with an affective valence (good or bad), but without any feeling of having gone through any steps of searching, weighing evidence or inferring a conclusion. (Greene and Haidt 2002: 517)¹⁶

¹⁵ According to Bayesians, if the temperature feels to us as if it is about 70 degrees, so we start with that assumption, then the fact that this thermometer reads 77 degrees should make us move our estimate towards 77 degrees. How much our estimate should increase depends on our prior assumption about how many thermometers work. This might make it seem as if the thermometer reading can lead to justified belief. However, if our initial estimates (about temperature and the percentage of working thermometers) are unjustified, then I doubt that one reading can ground justified belief. Besides, many of us do not form any of the initial estimates that are needed to start Bayesian reasoning.

¹⁶ See also Haidt 2001. Of course, many other judgments cause emotional reactions 'suddenly and effortlessly'. But Haidt argues that emotions drive or constitute moral judgments rather than being effects of those judgments.

Haidt's behavioral evidence dovetails nicely with independent brain studies. Moll's group found that brain tissue associated with emotions becomes more activated when subjects think about simple sentences with moral content (e.g. 'They hung an innocent') than when they think about similar sentences without moral content (e.g. 'Stones are made of water') (Moll *et al.* 2001) or disgusting non-moral sentences (e.g. 'He licked the dirty toilet') (Moll *et al.* 2002a).¹⁷ Similar results were found with pictures in place of sentences (Moll *et al.* 2002b).

Studies by Joshua Greene and his colleagues are even more fascinating because they distinguish kinds of moral beliefs (2001).¹⁸ Greene's group scanned brains of subjects while they considered what was appropriate in three kinds of dilemmas: non-moral dilemmas, personal moral dilemmas, and impersonal moral dilemmas. A moral dilemma is personal if and only if one of its options is likely to cause serious harm to a particular person other than by deflecting an existing threat onto a different party (Greene and Haidt 2002: 519). A standard personal moral dilemma is the fat-man trolley case. A paradigm impersonal moral dilemma is the side-track trolley case. These different moral cases stimulated different parts of the brain. While considering appropriate action in impersonal dilemmas, subjects showed significant activation in brain areas associated with working memory but no significant activation in areas associated with emotion. In contrast, while considering appropriate action in personal dilemmas, subjects showed significant activation in brain areas associated with emotion and under-activation (below the resting baseline) in areas associated with working memory. It is not obvious what to make of these results. Brain scientists do not know how to interpret under-activation in general. Nonetheless, one natural speculation is this: When asked about pushing the fat man, subjects react, 'That's so horrible that I can't even think about it.' Emotions stop subjects from considering the many factors in these examples. If this interpretation is correct, then many pervasive and fundamental moral beliefs result from emotions that cloud judgment.¹⁹

Some moral intuitionists might argue that there is no need to consider anything else when the proposed action is the intentional killing of an innocent fat man. It might even be counterproductive to consider additional factors, since they might lead one away from the correct belief. Such responses,

¹⁷ It would be interesting to test reactions to negations, such as 'They did not hang an innocent' and 'He did not lick the dirty toilet.'

¹⁸ This article also reports timing studies that confirm the different roles of emotion in different moral beliefs.

¹⁹ Philosophers should notice that what Greene calls 'personal dilemmas' include most proposed counterexamples to consequentialism. If those intuitions are unjustified, then Greene's study might help consequentialists defend their moral theory, even if other intuitions are not affected.

however, assume that it is morally wrong to push the fat man, so they beg the question here. When asking whether a moral belief is justified, we should not assume that the only relevant factors are those that would be relevant if the belief were true. Ridiculous moral beliefs could be defended if that method worked.

Still, moral intuitionism is hardly refuted by these experiments because Greene's results must be replicated and interpreted much more carefully. Our group at Dartmouth College is working on replications now, as is Greene. All I can say now is that such brain studies seem to provide some evidence that moral judgments result from emotions that cloud judgment.

Additional evidence comes from Wheatley and Haidt (2005). They gave participants the post-hypnotic suggestion that they would feel a pang of disgust whenever they saw either the word 'take' or the word 'often'. Participants were later asked to make moral judgments about six stories designed to elicit mild to moderate disgust. When a story contained the word that elicited disgust in a participant, that participant was more likely to express stronger moral condemnation of acts in the story. Moral judgments were then affected by elements of the story that could not determine the accuracy or acceptability of those moral judgments. In that sense, emotions clouded their judgment. Because independently caused emotions can distort moral beliefs in such ways, moral believers need confirmation in order to be justified in holding their moral beliefs.

Illusions

To apply Principle 4 to moral beliefs, we again need empirical research, but this time in cognitive science rather than brain science. I mentioned three kinds of illusions that should be considered separately.

The first kind of illusion occurs when appearances and beliefs depend on context. An interesting recent example comes from Peter Unger, who found that the order in which options are presented affects beliefs about whether a given option is morally wrong. He also claims that people's moral beliefs about a certain option depend on whether that option is presented as part of a pair or, instead, as part of a series that includes additional options intermediate between the original pair (Unger 1996: 88–94).²⁰ Since order and additional

²⁰ Unfortunately, Unger does not describe the method or precise results of his informal survey, so there is room for more careful empirical work to test his claims. Some philosophical support comes from moral paradoxes, which often arise through the mechanisms that Unger describes. One example is the mere addition paradox of Parfit 1984, in which B seems worse than A when the two are compared directly, but it seems not worse than A when Parfit interjects A + and Divided B as options intermediate between A and B.

options are not morally relevant factors that could affect the moral wrongness of the judged option, the fact that moral beliefs are affected by these factors shows that moral beliefs are unreliable in such cases. That is why confirmation is needed. One still might confirm one's moral belief by reconsidering the issue in several contexts over time to see whether one's moral belief remains stable, but that is just a way of seeking confirmation, so it does not undermine my point that confirmation is needed.

The second kind of illusion arises from overgeneralization. Such illusions also affect moral beliefs. Jonathan Baron even argues that all 'nonconsequentialist principles arise from overgeneralizing rules that are consistent with consequentialism in a limited set of cases' (1994: 1). But one need not accept consequentialism in order to admit that many people condemn defensible lying, harming, and love-making because they apply generalizations to exceptional cases. We probably disagree about which moral beliefs are overgeneralizations, but we should agree that many people overgeneralize in ways that create illusions of moral wrongness. In any such case, the moral believer could argue that this case is not an exception to the generalization, but, as before, that is just a way of seeking confirmation, so it does not undermine my point that this kind of illusion creates a need for confirmation.

Heuristics, which are quick and simple decision procedures, also create illusions in morality. One reason is that many moral beliefs depend on consequences and probabilities, for which we often lack adequate evidence, and then we have to guess these probabilities. Such guesses are notoriously distorted by the availability heuristic, the representative heuristic, and so on.²¹ Even when moral beliefs do not depend on probability assessments, moral beliefs are affected by the so-called 'I agree with people I like' heuristic (cf. Chaiken and Lord, Ross, and Lepper as discussed by Haidt 2001). When people whom we like express moral beliefs, we tend to go along and form the same belief. When people whom we dislike oppose our moral beliefs, we tend to hold on to them in spite of contrary arguments. This heuristic often works fine, but it fails in enough cases to create a need for confirmation.

In addition to these three kinds of illusions, moral beliefs also seem subject to framing effects, which were explored by Kahneman and Tversky (1979). In one famous experiment, they asked some subjects this question:

Imagine that the U.S. is preparing for an outbreak of an unusual Asian disease which is expected to kill 600 people. Two alternative programs to fight the disease, A and B, have been proposed. Assume that the exact scientific estimates of the consequences of

²¹ Kahneman *et al.* 1982. Lackey 1986: 634, discusses how such heuristics might explain conflicting moral intuitions about nuclear deterrence.

the programs are as follows: If program A is adopted, 200 people will be saved. If program B is adopted, there is a 1/3 probability that 600 people will be saved, and a 2/3 probability that no people will be saved. Which program would you choose?

The same story was told to a second group of subjects, but these subjects had to choose between these programs:

If program C is adopted, 400 people will die. If program D is adopted, there is a 1/3 probability that nobody will die and a 2/3 probability that 600 will die.

It should be obvious that programs A and C are equivalent, as are programs B and D. However, most subjects who chose between A and B favored A, but most subjects who chose between C and D favored D. More generally, subjects were risk averse when results were described in positive terms (such as 'lives saved') but risk seeking when results were described in negative terms (such as 'lives lost' or 'people who die').

The question in this experiment was about choices rather than moral wrongness. Still, the subjects were not told how the policies affect them personally, so their choices seem to result from beliefs about which program is morally right or wrong. If so, the subjects had different moral beliefs about programs A and C and about programs B and D. The only difference within each pair is how the programs are framed or described. Thus, descriptions seem to affect moral beliefs. Descriptions cannot affect what is really morally right or wrong. Hence, these results suggest that such moral beliefs are unreliable.

Moral intuitionists could claim that moral intuitions are still reliable when subjects have consistent beliefs after considering all relevant descriptions. But then moral believers would need to know that their beliefs are consistent and that they are aware of all relevant descriptions before they could be justified in holding moral beliefs. Framing effects distort moral beliefs in so many cases that moral believers need confirmation for any particular moral belief.

To see how deeply this point cuts, consider Warren Quinn's argument for the traditional doctrine of doing and allowing, which claims that stronger moral justification is needed for killing than for letting die. In support of this general doctrine, Quinn appeals to moral intuitions of specific cases:

In Rescue I, we can save either five people in danger of drowning at one place or a single person in danger of drowning somewhere else. We cannot save all six. In Rescue II, we can save the five only by driving over and thereby killing someone who (for an unspecified reason) is trapped on the road. If we do not undertake the rescue, the trapped person can later be freed. (1993: 152)

Most people judge that saving the five is morally wrong in Rescue II but not in Rescue I. Why do they react this way? Quinn assumes that these different

intuitions result from the difference between killing and letting die or, more generally, doing and allowing harm. However, Tamara Horowitz uses a different distinction (between gains and losses) and a different theory (prospect theory) to develop an alternative explanation of Quinn's moral intuitions:

In deciding whether to kill the person or leave the person alone, one thinks of the person's being alive as the status quo and chooses this as the neutral outcome. Killing the person is regarded as a negative deviation. . . . But in deciding to save a person who would otherwise die, the person being dead is the status quo and is selected as the neutral outcome. So saving the person is a positive deviation. . . . (1998: 153)

The point is that we tend to reject options that cause definite negative deviations from the status quo. That explains why subjects rejected program C but did not reject program A in the Asian disease case (despite the equivalence between those programs). It also explains why we think that it is morally wrong to 'kill' in Rescue II but is not morally wrong to 'not save' in Rescue I, since killing causes a definite negative deviation from the status quo. This explanation clearly hinges on what is taken to be the status quo, which in turn depends on how the options are described. Quinn's story about Rescue I describes the people as already 'in danger of drowning', whereas the trapped person in Rescue II can 'later be freed' if not for our 'killing' him. These descriptions affect our choice of the neutral starting point. As in the Asian disease cases, our choice of the neutral starting point then affects our moral intuitions. Horowitz adds, 'I do not see why anyone would think the distinction [that explains our reactions to Quinn's rescue cases] is morally significant, but perhaps there is some argument I have not thought of. If the distinction is not morally significant, then Quinn's thought experiments do not support one moral theory over against another' (1998: 155).

Admittedly, Horowitz's explanation does not imply that Quinn's moral intuitions are false or incoherent, as in the Asian disease case. It does not even establish that his moral intuitions are arbitrary. As Mark van Roojen says, 'Nothing in the example shows anything wrong with treating losses from a neutral baseline differently from gains. Such reasoning might well be appropriate where framing proceeds in a reasonable manner' (1999).²² Nonetheless, the framing also 'might well' *not* be reasonable, so the epistemological dilemma remains: If there is no reason to choose one baseline over the other, then our moral intuitions seem arbitrary and unjustified. If there is a reason to choose one baseline over the other, then either we have access to that reason or we do not. If we have access to the reason, then we are able to draw an inference from

²² Van Roojen might admit that Horowitz's argument undermines moral intuitionism, since he defends a method of reflective equilibrium that is coherentist rather than foundationalist.

that reason to justify our moral belief. If we do not have access to that reason, then we do not seem justified in our moral belief. Because framing effects so often lead to incoherence and error, we cannot be justified in trusting a moral intuition that relies on framing effects unless we at least can be aware that this intuition is one where the baseline is reasonable. So Horowitz's explanation creates serious trouble for moral intuitionism whenever framing effects could explain our moral intuitions.

The doctrine of doing and allowing is not an isolated case. It affects many prominent issues and is strongly believed by many philosophers and common people, who do not seem to be able to infer it from any other beliefs. If moral intuitions are unjustified in this case, doubts should arise about a wide range of other moral intuitions as well.

Origins

Some previous principles look at origins of individual moral beliefs, but Principle 5 considers the social origins of shared moral beliefs. The two issues are related insofar as many of our moral beliefs result from training and social interaction.

Specifically, Principle 5 claims that problematic social origins create a need for confirmation. To apply this principle, we need to ask whether moral beliefs have problematic social origins. The social origins of moral beliefs might be problematic in two ways. First, moral beliefs might be caused by factors that are unrelated with the truth of those beliefs. Second, the origins of moral beliefs might be immoral according to those moral beliefs. I will focus on the latter case.

Are the origins of our moral intuitions immoral by their own lights? Friedrich Nietzsche suggests as much when he argues that Christian morality results from slaves cleverly overcoming their superiors by re-evaluating values. Insofar as Christian morality condemns such subterfuge and self-promotion, Christian morality condemns its own origins, if Nietzsche is correct (Nietzsche 1966).²³ Similarly, Michel Foucault argues at length that moral beliefs express or result from social power relations. Yet these moral beliefs themselves seem to condemn the very kind of power that leads to these beliefs. But I don't want to rely on Nietzsche or Foucault, at least not in this crowd, so I will consider Gilbert Harman's explanation of the common moral belief that harming someone is much worse than failing to helping someone in need:

whereas everyone would benefit equally from a conventional practice of trying not to harm each other, some people would benefit considerably more than others from a

²³ I am not, of course, endorsing Nietzsche's speculations.

convention to help those who needed help. The rich and powerful do not need much help and are often in the best position to give it; so, if a strong principle of mutual aid were adopted, they would gain little and lose a great deal, because they would end up doing most of the helping and would receive little in return. On the other hand, the poor and the weak might refuse to agree to a principle of non-interference or non-injury unless they also reached some agreement on mutual aid. We would therefore expect a compromise [that] would involve a strong principle of non-injury and a weaker principle of mutual aid—which is just what we now have. (1977: 110; cf. Scheffler 1982: 113)

Remember also that rich and powerful people have always controlled the church, the media, and culture, which in turn affect most people's moral beliefs. In this context, Harman's claim is that the self-interest of the rich and powerful in making everyone believe that harming is worse than failing to help can explain why so many people believe that harming is worse than failing to help. But our moral beliefs also seem to condemn such self-serving indoctrination by the rich and powerful, since morality is supposed to consider everyone's interests equally. Thus, if Harman is correct, morality condemns its own origins, as Nietzsche and Foucault claimed.

The point is not that such moral views are internally inconsistent, selfcondemning, or even self-defeating. The point is only that there are grounds for doubt when beliefs come from disreputable sources. Defenders of such moral beliefs must admit that the sources of their beliefs are disreputable if Harman's explanation is accurate. Then they need additional support for their beliefs beyond the mere fact that those beliefs seem correct to them.

These speculations about the origins of moral beliefs are mere armchair psychology. Perhaps more support could be obtained from the literature on sociobiology or evolutionary psychology. Still, these explanations are likely to remain very controversial. Luckily, I don't need to prove them here. I claim only that these undermining accounts are live possibilities. They seem plausible to many people and have not been refuted.

That would not be enough if I were arguing for the falsehood of a certain moral belief, such as Christian morality (from Nietzsche) or the prevalence of non-injury over mutual aid (from Harman). However, I am not drawing any substantive moral conclusion. To do so would commit a genetic fallacy, but my argument is different. My point lies in moral epistemology, and I reach it indirectly. If these disreputable origins are live possibilities, then moral believers need some independent confirmation that their beliefs are not distorted by such disreputable origins. This need for independent confirmation then undermines moral intuitionism.

Togetherhness

Don't forget that Principles 1–5 complement each other. If I am right, moral beliefs are partial, controversial, emotional, subject to illusion, and explicable by dubious sources, so all of the principles apply. However, even if not all but only several of them apply, these principles still work together to make it clear that confirmation is needed for justified moral belief. That undermines moral intuitionism. It also shows how empirical research can be indirectly relevant to normative moral epistemology.

4. Objections

None of my arguments is conclusive, so opponents can object at several points. Here I cannot respond to every objection or to any objection thoroughly. But I will quickly run through the most formidable objections.

Confirmation

One common objection is that, even if some confirmation is needed, that does not show that any inference is needed. If we can confirm color beliefs just by looking again in different light, perhaps we can confirm moral beliefs simply by reflecting on the moral issue again in a different mood without involving any substantive moral principle from which we infer our moral belief.

I grant that confirmation does not require an actual inference. To avoid the skeptical regress, however, moral intuitionists must deny more than the need for an actual inference. They must deny the need for any ability to infer the moral belief. It is hard to see how you could confirm a moral belief without gaining information that makes you able to draw some kind of inference to the moral belief. Even if you just think about the moral issue several times in different moods, after such rethinking you have all the information you need for a simple inference like this:

I hold this moral belief after reflecting on the issue several times in different moods.

If I hold a moral belief after reflecting on the issue several times in different moods, then it is usually true.

So, probably, this moral belief is true.

Admittedly, this inference is not a deductive proof. Nor does it infer the moral belief from a more general substantive moral principle. But no specific kind of

inference is needed. Any kind of inference can lead to a skeptical regress, so moral intuitionists have to deny dependence on any kind of inference or ability to infer. And an ability to draw the above kind of inference is needed, since, if the moral believer does not believe its premises (or something like them), then it is hard to see why the moral believer is justified in holding the moral belief.

At this point, externalists (including reliabilists) sometimes accuse me of confusing whether a belief is justified with whether the believer knows that it is justified. I plead innocent. I do not assume that justified believers must know or be justified in believing (or even be able to know or be justified in believing) that they are justified. I claim only that justified moral believers must be able to infer their moral beliefs from something. Some externalists still deny this, but their denial is implausible, as I have argued elsewhere (Sinnott-Armstrong 2002*b*). Besides, most externalists are mainly concerned about non-moral beliefs, including perceptual beliefs. Principles 1–5 do not apply to perceptual beliefs in the same way as to moral beliefs. Most perceptual beliefs are not partial, controversial, emotional, or explicable by dubious sources. Perceptual illusions are common, but they are normally easier to detect than moral illusions, and they do not affect anything as widespread and fundamental as the doctrine of doing and allowing. Consequently, externalism and reliabilism might work for perceptual beliefs even if, as I have argued, moral beliefs need confirmation of a kind that requires an inferential ability.²⁴

Children

This response leads to another objection. It might seem too strict to require an inferential ability because then young children cannot have justified moral beliefs, since they cannot formulate the needed inferences.

I love children. I grant that they can have justified beliefs in other areas, such as beliefs about food and toys. However, it is not as clear that very young children (say, 1–3 years old) can be justified in holding moral beliefs. After all, young children often base their normative beliefs on fear of punishment. If someone believes that stealing is wrong just because he believes that he is likely to get punished if he steals, then it is not even clear that the belief is a moral belief. It might be purely prudential. Instead of fear, the basis for some young children's moral beliefs might be deference to authorities (or peers). But if children accept their parents' word that an act is wrong without any idea of what makes that act wrong, then these children might not believe that the act is wrong *morally*, since they might not believe that there is any specifically

²⁴ I do not claim that moral judgments are the only ones that need confirmation according to Principles 1–5. Several of these principles also apply to beliefs about what is prudent, rational, wise, and beautiful.

moral reason not to do the act. Moreover, authorities can make someone justified only if she is justified in trusting those authorities. Maybe young children are justified in trusting their parents, for example. But then they can infer:

My parents are trustworthy.

My parents tell me that I shouldn't pull my sister's hair.

Therefore, I shouldn't pull my sister's hair.

If a young child is not able to draw any inference like this, then this child does not seem justified (even if her parents are trustworthy and even if her belief is true). Those who think otherwise are too soft on their kids.

Ignorance

Another objection claims that, if a moral believer could not know that moral beliefs are subject to controversy, partiality, illusion, and so on, then that moral believer does not need to guard against these problems by getting confirmation. Children and some adults (such as isolated medieval peasants) might have no way of discovering such problems for moral beliefs. They certainly lack access to the psychological research that I cited. So maybe these moral believers do not need confirmation for their moral beliefs.

This objection confuses two claims. To call a believer unjustified is often to criticize that believer. Such criticism seems misplaced when the believer is not responsible for any epistemic failures. Believers are not responsible when they have no way of knowing that their beliefs are problematic. Thus, if children and medieval peasants cannot know that moral beliefs are problematic, it seems odd to call them unjustified.

In contrast, to say that a believer is not justified is not to criticize the believer. It is only to withhold the praise of calling the believer justified. There is nothing unfair about withholding praise when a believer is not responsible. Thus, even if children and medieval peasants are not responsible for their epistemic failures, that does not undermine my claim that they are not justified in their moral beliefs.

Moreover, even if children and medieval peasants were justified in their moral beliefs, that would not save moral intuitionists or my readers from the need for confirmation. Moral intuitionists and my readers are neither children nor medieval peasants. They are modern educated adults. Modern educated adults can know that moral beliefs are problematic in the ways that I outlined (at least if they have read this far). So my readers and other modern adults need confirmation for moral beliefs, regardless of what you think about other people.

Moral intuitionists might seem to avoid this point if they claim only that some moral believers are justified. However, moral intuitionists always include themselves among those who are justified. Similarly, I assume that my readers want to know whether they themselves are justified moral believers. If it turns out that the only moral believers who are justified without confirmation are children, medieval peasants, and others who are ignorant of the empirical research in this paper, then it is not so great to be justified.

Defeasibility

Some moral intuitionists accuse me of forgetting that a moral believer can be defeasibly justified without being adequately justified. Again, I plead innocent.

To say that a moral believer is defeasibly justified is to say that the believer would be justified in the absence of any defeater. Defeaters come in two kinds. An overriding defeater of a belief provides a reason to believe that the belief is false. For example, if one newspaper predicts rain tomorrow, but a more reliable newspaper predicts clear skies, then the latter prediction overrides the former, even if I still have some reason to believe the former. In contrast, an undermining defeater takes the force out of a reason without providing any reason to believe the opposite. If I find out that the newspaper that predicts rain based its prediction on a crystal ball, then this new information keeps the prediction from making me justified in believing that it will rain, but the new information does not make me justified in believing that it will *not* rain, since a crystal ball is just as likely to lead to a true prediction. When my justification is undermined completely in this way, I have no reason left for believing that it will rain or that it will not rain.

The factors in Principles 1–5 cannot be overriding defeaters, since they do not provide any reason to believe the moral belief is false. Even when moral beliefs are partial, controversial, emotional, subject to illusion, and due to disreputable sources, that does not show that those beliefs are false. Thus, the factors in Principles 1–5 seem to be undermining defeaters. That suggests that we have no reason to trust our spontaneous moral beliefs before confirmation. Admittedly, some defeaters might not completely undermine a justification. They might leave some weaker reason that makes believers partially justified. However, the manifold underminers in Principles 1–5 add up, so that it is hard to see why there is any reason left to hold spontaneous moral beliefs without confirmation.

Moreover, I am not just talking about *possible* underminers. I argued in section 3 that the underminers in Principles 1–5 actually exist for many moral

beliefs. Actual moral believers are partial and emotional. They actually do disagree often. Cultures actually are disreputable in ways that affect moral beliefs. There is even empirical evidence for actual widespread illusions in morality.

Moral intuitionists can still say that spontaneous moral believers are *prima facie* justified if that means only that they would be adequately justified if their moral beliefs were not undermined by the factors in Principles 1–5. This counterfactual claim is compatible with their actually not being justified at all, but only appearing to be justified. They might have no real reason for belief but only the misleading appearance of a reason (as with the newspaper's rain prediction based on a crystal ball). In contrast, to call a believer *pro tanto* justified is to indicate some actual positive epistemic force that is not cancelled or undermined even if it is overridden. If the factors in Principles 1–5 are underminers, as I argued, then spontaneous moral believers are not even *pro tanto* justified. At most they misleadingly appear to be justified when they are not really justified at all.

Besides, even if moral intuitions were *pro tanto* justified independently of any inferential ability, this status would not make them adequately justified. As I said, skeptics win if no moral belief is adequately justified. So moral intuitionists cannot rest easy with the claim that moral intuitions are merely *pro tanto* justified.

Some

Many opponents object that, even if Principles 1–5 apply to some moral beliefs, they do not apply to all moral beliefs. As I admitted, some moral beliefs are not controversial. For example, almost everyone (except moral nihilists) agrees that it is morally wrong to push the fat man in front of the trolley just because you are angry with him for beating you in a game. Such cases also do not seem due to context, heuristics, overgeneralization, or framing effects. Still, such moral believers are partial and emotional (as Greene's experiments suggested). So Principles 1 and 3 do seem to create a need for confirmation even in such clear cases.

Furthermore, if very many moral beliefs need confirmation, the others cannot be immune from this need. To see why, compare a country with lots of barn façades that look just like real barns when viewed from the road (Goldman 1976). If someone looks only from the road, then he is not justified in believing that what he sees is a real barn, at least if he should know about the barn façades. The barn façades are analogous to situations that produce distorted moral beliefs. Since such distortions are so common, morality is a land of fake barns. In such areas, confirmation is needed for each justified

belief, even for those beliefs formed in front of real barns. Analogously, confirmation is needed for each spontaneous moral belief, even when the common distorting factors are absent. We need to get off the road and look closer. At least when we should know that moral beliefs in general are so often subject to distortion, we cannot be justified in trusting any moral belief until we confirm that it is an exception to the rule that most moral beliefs are problematic. So moral intuitionists cannot claim that any moral believers are justified without confirmation.

This point can be presented as a dilemma: If a moral believer is an educated modern adult, then she should know that many moral beliefs are problematic in the ways indicated by Principles 1–5. She either knows or does not know that her moral belief is an exception to the trend. If she does not know this, she should accept a significant probability that her belief is problematic. Then she cannot be justified without confirmation. Alternatively, if she does know that her moral belief is exceptionally reliable, then she has enough information to draw an inference like this: My moral belief is exceptionally reliable. Exceptionally reliable beliefs are probably true. Therefore, my belief is (probably) true. If this moral believer does not have the information in these premises, then it is hard to see why we should call her justified. So, either way, moral intuitionism fails.

Skepticism

A common objection is that my argument leads to general skepticism, since every inference has premises, so the demand for an inference cannot always be met. However, my argument does not generalize so easily. If my belief that a pen is in front of me is not subject to disagreement or illusions and has no disreputable sources, and if I am neither partial nor emotional about pens, then I might be justified in holding that non-moral belief without being able to support it with any inference. Thus, my argument against moral intuitionism does not lead to general skepticism.

My argument still might seem to lead to moral skepticism. If so, and if moral skepticism is unacceptable, then something must be wrong with my argument. However, my argument does not by itself lead to moral skepticism. My thesis is not that spontaneous moral beliefs are not justified, but only that they are not justified non-inferentially because they need confirmation. Such confirmation still might be possible somehow. Even if moral intuitionism is rejected, there are other non-skeptical methods in moral epistemology, including coherentism, contractarianism, contractualism, contextualism, and naturalism (Sinnott-Armstrong 1996: 31–41). Moral skepticism arises only after all of these other approaches fall. So my argument does not by itself support moral skepticism.

Besides, even if these other approaches also fail, so my argument plays a role in a larger argument for moral skepticism, that does not show that anything is wrong with my argument, unless one assumes that moral skepticism is unacceptable. Why assume that? I accept a limited Pyrrhonian version of moral skepticism. So I, at least, will not be dismayed if my argument takes one step in that direction.²⁵

Anyway, my goal here has not been to argue for moral skepticism. My goal has been to argue against moral intuitionism. More generally, I tried to show one way in which empirical research in psychology and brain science might be relevant to normative moral epistemology. If I succeeded in that enterprise, I will happily leave moral skepticism for another occasion.²⁶

References

- BARON, J. (1994). 'Nonconsequentialist Decisions'. *Behavioral and Brain Sciences*, 17: 1–42.
- FOOT, PHILIPPA (1967). 'The Problem of Abortion and the Doctrine of Double Effect'. *Oxford Review*, 5: 5–15.
- GOLDMAN, A. (1976). 'Discrimination and Perceptual Knowledge'. *Journal of Philosophy*, 73: 771–91.
- GREENE, J., *et al.* (2001). 'An fMRI Investigation of Emotional Engagement in Moral Judgment'. *Science*, 293: 2105–8.
- and HAIDT, J. (2002). 'How (and Where) does Moral Judgment Work?' *Trends in Cognitive Science*, 6: 517–23.
- HAIDT, J. (2001). 'The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment'. *Psychological Review*, 108: 814–34.
- *et al.* (1993). 'Affect, Culture, and Morality, or is it Wrong to Eat Your Dog?' *Journal of Personality and Social Psychology*, 65: 613–28.
- HARMAN, G. (1977). *The Nature of Morality*. New York: Oxford University Press.
- HOROWITZ, T. (1998). 'Philosophical Intuitions and Psychological Theory'. In M. DePaul and W. Ramsey (eds.), *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*. Lanham, Md.: Rowman & Littlefield.
- KAHNEMAN, D., SLOVIC, P., and TVERSKY, A. (eds.) (1982). *Judgment under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

²⁵ My version of Pyrrhonism denies that any particular contrast class is the relevant one in the sense that believers need to rule out all alternatives in that class in order to be justified without qualification. That view might seem to conflict with my claim here that moral beliefs need inferential confirmation. However, there might be no need to rule out any particular contrast class, even if there is a need to give evidence of a certain kind.

- and TVERSKY, A. (1979). 'Prospect Theory: An Analysis of Decision under Risk'. *Econometrica*, 47/2: 263–92.
- LACKEY, D. (1986). 'Taking Risk Seriously'. *Journal of Philosophy*, 83: 633–40.
- MOLL, J., *et al.* (2001). 'Frontopolar and Anterior Temporal Cortex Activation in a Moral Judgment Task: Preliminary Functional MRI results in Normal Subjects'. *Arq. Neuropsiquiatr*, 59: 657–64.
- *et al.* (2002a). 'Functional Networks in Emotional Moral and Nonmoral Social Judgments'. *Neuroimage*, 16: 696–703.
- *et al.* (2002b). 'The Neural Correlates of Moral Sensitivity: A Functional Magnetic Resonance Imaging Investigation of Basic and Moral Emotions'. *Journal of Neuroscience*, 22: 2730–6.
- MOORE, G. E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.
- NIETZSCHE, F. (1966). *Genealogy of Morals*. In *Basic Writings of Nietzsche*, ed. and trans. W. Kaufmann. New York: Random.
- PARFIT, D. (1984). *Reasons and Persons*. Oxford: Clarendon Press.
- QUINN, W. (1993). 'Actions, Intentions, and Consequences: The Doctrine of Doing and Allowing', repr. in *Morality and Action*. New York: Cambridge University Press.
- ROSS, W. D. (1930). *The Right and the Good*. Oxford: Oxford University Press.
- SCHEFFLER, S. (1982). *The Rejection of Consequentialism*. Oxford: Clarendon Press.
- SINNOTT-ARMSTRONG, W. (1996). 'Moral Skepticism and Justification'. In W. Sinnott-Armstrong and M. Timmons (eds.), *Moral Knowledge?* New York: Oxford University Press.
- (2000). 'From "Is" to "Ought" in Moral Epistemology'. *Argumentation*, 14: 159–74.
- (2002a). 'Moral Skepticism', in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2002 edn.), <http://plato.stanford.edu/archives/sum2002/entries/skepticism-moral/>.
- (2002b). 'Moral Relativity and Intuitionism'. *Philosophical Issues*, Realism and Relativism, 12: 305–28.
- SNARE, F. (1980). 'The Diversity of Morals'. *Mind*, 89: 353–69.
- UNGER, P. (1996). *Living High and Letting Die*. New York: Oxford University Press.
- VAN ROOJEN, M. (1999). 'Reflective Moral Equilibrium and Psychological Theory'. *Ethics*, 109: 846–57.
- WHEATLEY, T., and HAIDT, J. (2005). 'The Wisdom of Repugnance: Hypnotically-Induced Disgust Makes Moral Judgments More Severe' (unpublished manuscript, University of Virginia).

