Intuitional Epistemology in Ethics
Matthew S. Bedke*
UBC Philosophy

Abstract
Here I examine the major theories of ethical intuitions, focusing on the epistemic status of this class of intuitions. We cover self-evidence theory, seeming-state theory, and some of the recent contributions from experimental philosophy.

1. Our Topic
It is hard to over-exaggerate the role ethical intuitions play in contemporary ethical theorizing. Just consider the following sample of intuitions we have been invited to share:

1. An intuition about a footbridge version of a trolley case: that it is impermissible to push a fat man off a footbridge to stop a train bearing down on five people, killing the fat man, but saving five equally innocent others (Thomson 1985).
2. Contrasting intuitions about doing-allowing cases, e.g., that it is impermissible to kill one to save five, while it is permissible to let one die to save five (Quinn, 1989).
3. An intuition about the repugnant conclusion: that a world with many people all with utilities slightly above the zero point is not better than a world with fewer people with utilities significantly above the zero point, even if the first world features more aggregate utility than the second world (Parfit 1984: Chapter 17).
4. An intuition that one has a *prima facie* duty to fulfill one’s promises (alternatively put, that one has promised to φ is a (moral) reason to φ) (Ross 1930: Chapter 2).
5. An intuition about Sidgwick’s axiom of the Universal Good: that ‘The good of any one individual is of no more importance, from the point of view, if I may say so, of the Universe, than the good of any other’ (Sidgwick 1907: 382).
6. An intuition about the transitivity of value: that if a is better than b, and b is better than c, then a is better than c.

In each case, the thought is that having the proffered intuition helps to justify the corresponding belief, where the justification is epistemically non-inferential. That is, the justification does not result from an inference that transmits justification from some premises to the moral conclusion.

Let us then state the core idea of epistemological ethical intuitionism like this: if A has an intuition that p and believes that p on the basis of that intuition (where the basing relation is understood non-inferentially), A is to some degree justified in believing that p (unless or until A acquires defeaters). To put things in terms of propositional justification, one might say that an intuition that p is a defeasible epistemic reason to believe that p.1

Here we consider the nature of ethical intuitions and their capacity to non-inferentially justify ethical beliefs. Our guiding questions are these: What are ethical intuitions? And what is the correct epistemological theory of how they non-inferentially justify beliefs (if they do so)? In so framing the questions we will largely presuppose that the ethical judg-
ments supported by intuition are cognitive, belief-like, mental states. But we should take care to distinguish our topic from Ethical Intuitionism as traditionally conceived. The latter covers more ground, including not only an intuitional epistemology and a cognitivist moral psychology, but three other views as well: (i) the correct moral theory is a set of mid-level principles identifying those features that make right actions right and wrong actions wrong; (ii) the principles are pluralist in identifying more than one feature in virtue of which right actions are right and wrong actions wrong; and (iii) ethical properties are metaphysically real and irreducible to non-ethical properties. We will only consider these other positions if and when prompted to do so by our primary interest in the nature and epistemology of ethical intuitions.

Let me proceed by distinguishing two influential answers to our guiding questions: self-evidence theory (section 2) and seeming state theory (section 3). Along the way we will have occasion to consider some of the more pressing issues in the contemporary debates, such as whether ethical intuitions are of a kind with other intuitions in philosophy. The burgeoning area of experimental philosophy and its impact on intuitional epistemology in ethics will be addressed in section 4.

2. Self-Evidence Theory

Historically, most philosophers who advocate an intuitional moral epistemology speak of the contents of justified ethical beliefs/knowledge in terms of self-evident propositions or facts. Exactly what is taken to be self-evident varies from philosopher to philosopher. W. D. Ross, the quintessential ethical intuitionist, thought that certain mid-level principles of prima facie duty were self-evident, e.g., the promise-keeping duty of fidelity in example 4 above. For him, the self-evidence of moral truths is analogous to the self-evidence of mathematical axioms. His views led him to embrace a deontological moral theory that was pluralist about right- and wrong-making features of actions.

Sidgwick, by contrast, would disagree about which propositions are self-evident, largely because a pluralist view like Ross’s cries out for some rational synthesis, and more determinate guidance on all-in questions about what to do. He proposed roughly the following as self-evident: that the good of any one person is as important as the good of any other (see example 5 above), and that one is rationally bound to aim at the good generally, not any part of it (Sidgwick 1907: 382). From these and supplementary claims about the good, he argues for a form of consequentialism called Hedonic Utilitarianism.

Sidgwick had more to say about what it is for a proposition to be self-evident, or at least the circumstances under which a proposition can play the epistemic role we think self-evident propositions play. In his words, he sets forth four conditions sufficient to establish ‘a significant proposition, apparently self-evident, in the highest degree of certainty attainable’: (i) the target proposition must be clear and precise; (ii) it’s self-evidence must be ascertained by careful reflection; (iii) taken as a whole, propositions taken as self-evident must be mutually consistent; and (iv) the target proposition must lack universal or general denial by others, or at least lack denial among experts (Sidgwick 1907: 211–12, 338–42). Realizing that others disagreed with him about what propositions were self-evident, he also warned that it is easy to confuse intuitions with blind impulses, vague sentiments, preferences for action, conclusions from rapid half-conscious processing, and current opinions to which familiarity has given an illusory air of self-evidence, such as conventional law (Sidgwick 1907: 211–12).

Despite their differences, both Ross and Sidgwick looked to principles or propositions of some generality for self-evident truth. Others have sought self-evidence in propositions
about concrete, particular actions, either actual or hypothetical. Both Samuel Clarke and H. A. Prichard can be read in this way (Clarke 1706: 226 (academic pagination); Prichard 1912: 28). Faced with moral decisions in real-life cases, individuals can just see that certain responses are fitting, or that certain actions are morally obligatory, or so the thought goes. In this vein, it is noteworthy that Ross thought that what came first in time was the self-evidence of prima facie duties to do certain concrete, particular actions, from which we intuitively induce the more general principles (Ross 1930: 33; 1939: Ch. 8).

To evaluate these claims, we need to know more about two things. First, we need to have a better idea of the epistemic role self-evident propositions supposedly play and so a better idea of the epistemic properties of corresponding intuitions. Second, we need to know more about the non-epistemic nature of self-evident propositions and corresponding intuitions such that they would have the claimed epistemic roles or properties.

Contemporary self-evidence theory can be seen in the first instance as an attempt to address the first point. That is, it is fair to say that self-evidence theory begins by characterizing the epistemic role of a certain class of propositions. Robert Audi, for instance, suggests that a self-evident proposition is ‘a truth’ such that ‘an adequate understanding of it is sufficient both for being justified in believing it and for knowing it if one believes it on the basis of that understanding’ (Audi 1998: 20, 1999: 206, 2004: 49; see also Audi 2008: 478). Similarly, Russ Safer-Landau says ‘A proposition $p$ is self-evident = df. $p$ is such that adequately understanding and attentively considering just $p$ is sufficient to justify believing that $p$’ (Shafer-Landau 2003: 247). Both authors further clarify these epistemic roles as follows: self-evident justification does not entail indefeasibility, epistemic support for intuitively justified beliefs can be strengthened or weakened via non-intuitive modes of justification, intuitive justification does not require that one see a proposition’s self-evidence, or have any intuitively justified beliefs about a proposition’s self-evidence, and one can adequately understand a self-evident proposition and yet fail to believe it (see, e.g., Audi 2004: Ch. 2, and Shafer–Landau 2003: Ch. 11).

With some understanding of self-evident propositions along these lines, one can then define intuitions in terms of self-evident propositions. Audi suggests something like this: (some) ethical intuitions are dispositions to believe self-evident ethical propositions based on non-inferential impressions of their truth, where the impression of truth flows from adequate understanding (cf. Audi 2008: 477–78).

This work is crucial when it comes to clarifying the epistemic role of intuitions in ethical theorizing, and particularly what epistemic theses intuitionists are committed to. More importantly, it allows us to consider with a clear view whether there are propositions that play this role. Any adequate answer to this question will have to say something about the non-epistemic nature of putatively self-evident propositions that explains why they enjoy such epistemic favor. After all, not any old proposition adequately understood can justify belief in it. Because there is significant consensus on the epistemic roles now claimed, in the remainder of this section I will focus on this explanatory task.

### 2.1. THE CONTAINMENT VIEW

Audi lately suggests that a self-evident proposition is one where the concept of the subject contains the concept of the predicate (Audi 2008: 479). Conceptual containment is typically associated with analytic propositions, where one tries to explain what it is for a proposition to be true in virtue of the meanings of its constituents. Thus, ‘a vixen is a female’ is analytically true (or so the proposal goes) because the concept of being a vixen contains the concept of being a female.
Audi suggests, however, that not all cases of containment are associated with analytic propositions. A concept C-1 might contain another concept C-2 even though there is no full analysis of C-1 in terms of the concepts it contains. In support of this idea, Audi proposes the following example: that burning a person with a red-hot poker would cause him excruciating pain is a moral reason not to do it (Audi 2008: 479). Here, the suggestion goes, the concept MORAL REASON FOR ACTION contains the concept AVOIDANCE OF CAUSING PAIN, though it takes adequate understanding of these concepts and some reflection to see this fact.

One interesting thing about containment view is that it would allow one to explain the epistemic properties of self-evident propositions in terms of semantics and concept possession. If a proposition simply reflects the containment relations of the concepts expressed, and we have some access to our concepts and their relations, it is easy to see why self-evident propositions are true, and why they strike us as true.³

Be that as it may, are there any special problems with the containment story as applied to substantive ethical propositions? Consider again the proposal that the concept FEMALE is contained within the concept VIXEN. If this is so, then thinking a thought about vixenhood contains (at least implicitly) a thought of femaleness, so that to be categorizable under VIXEN an item must be categorizable under FEMALE. The important point for our purposes is that a claim like ‘a vixen is a female’ falls out of the containment view as a conceptual necessity. This would be so even if the concept VIXEN did not admit of a full analysis or a definition in terms of necessary and sufficient conditions.

Now consider Audi’s example. If MORAL REASON FOR ACTION contains AVOIDANCE OF CAUSING PAIN, should we not expect ‘an action supported by a moral reason is an action that avoids the causing of pain’ to be conceptually necessary, even if MORAL REASON FOR ACTION does not admit a full analysis or definition? It would seem so if the vixen example is a good guide. This would be a decisive objection to the view, for there are a number of conceivable actions supported my moral reasons that have nothing to do with pain avoidance.

Audi’s view is likely a variation on the theme,⁴ where MORAL REASON FOR ACTION contains the thought that avoidance of causing pain is among the actions supported by moral reasons. In other words, the concept MORAL REASON FOR ACTION contains the concept PAIN AVOIDANCE BEING SUFFICIENT TO BE SUPPORTED BY MORAL REASON. One remaining worry is that this generates conceptually necessary truths where there are none, for it looks as though ‘an action that avoids the causing of pain is an action supported by a moral reason’ would turn out to be a conceptual necessity on this containment view. While some might think substantive ethical propositions like this can be conceptually necessary truths, many doubt it. The more common view is that some such substantive propositions are necessarily true, but not conceptual necessities (i.e. a kind of a synthetic necessity).

2.2. APPLICATION INTUITIONS

Perhaps a larger concern is that the containment view now has a concept containing itself, or at least referring to itself. Talk of containment might be getting in the way at this point. The crucial point for Audi is that certain non-moral facts ground the applicability of moral concepts (Audi 2008: 479). That an action would avoid the causing of pain would ground application of the concept MORAL REASON FOR ACTION is the key thought. So I now want to talk about another characterization of propositions that are self-evident that is also meant to help explain their epistemic properties and that
highlights the following idea: a self-evident proposition predicates a concept of conditions sufficient for the concept to apply. This is a view that might subsume ethical intuitions as instances of intuitions in other fields of philosophy, so let me begin by considering a plausible theory of what (some) philosophical intuitions are and how they justify.

Consider a Gettier-type case where Jones looks at a clock that reads 12:30 PM, and on that basis he believes truly that it is 12:30 PM. Unknown to him, the clock is broken and stuck on 12:30 PM. Plausibly, when asked whether Jones knows that it is 12:30 PM, one’s intuition that he does not know just is, or is a manifestation of, a disposition to withhold application of the concept KNOWLEDGE from the case driven by one’s competent know-how with the semantic norms governing application of KNOWLEDGE. Why semantic norms? One conspicuous feature of concept application in a case like this is its modal force. For those who have the intuition it seems like Jones’s belief just can’t be knowledge. Not all dispositions to apply concepts have this feature, which invites the hypothesis that concept application here is being driven by semantic-value fixing norms, as opposed to pragmatic, epistemic or other norms that might dispose one to treat cases one way or another, and which cannot determinately rule out a given case as non-knowledge.5

Following up on this idea, we might say that for \( p \) to be self-evident to \( A \) is for \( p \) to be a proposition that results from \( A \)’s disposition to apply (withhold) a concept to a case sufficient for the concept to apply, where the disposition is driven by semantic norms. And \( A \)’s intuition that \( p \) is a sense of \( p \)’s truth that results from \( A \)’s semantically-driven disposition to apply (withhold) the concept. This yields another semantic explanation for the justificatory properties of self-evident propositions, for any belief that is based on the semantic-value fixing norms governing concept application is likely true, assuming otherwise favorable epistemic circumstances. Here, it is important that we are only talking about norms that fix semantic value. Not all norms governing concept application would generate beliefs that are likely true. It is also important that we not take it for granted that certain instances of concept application are driven by semantic norms, for there are large outstanding questions about how semantic values are fixed, and so whether there are semantic norms that do the fixing and that are accessible to cognition in a way that can drive concept application.

As suggested above, proponents of the semantic norm explanation should take the fact that we have intuitions of a special modal character as \( data \) best explained by positing the accessibility and deployment of semantic-value fixing norms in these cases. If someone tells me that Jones does not know it is 12:30 PM, and I take that person to be credible, I might on that basis think that Jones doesn’t know it is 12:30 PM. Would I think that this can’t be a case of knowledge? Of course not. This classification of non-knowledge is very different from the classification based on the Gettier-type description. Why these classifications would be different, particularly in modal force, cries out for explanation. The suggestion is that the best explanation for the modal force of certain classifications posits the deployment of semantic norms governing concept application, norms that are not deployed in cases of concept application that lack this modal force.

In any event, let us call intuitions that fit this general model \( application intuitions \).6 Plausibly, many philosophically interesting intuitions are application intuitions, or have the semblance of application intuitions. Are some ethical intuitions application intuitions?

Though, to my knowledge, no one in the ethics literature has put things in exactly these terms, I think it aptly captures views implicit in the way proponents of intuitional epistemology proceed. Consider the Footbridge case, averred to in example 1 above. There we are presented with a hypothetical case, that of pushing an innocent fat man off
a footbridge as the only way to stop a runaway trolley from killing five other innocent people. Many are inclined to classify the action as impermissible and have some impression that 'it is impermissible to kill the one to save the five in this case' is true. Moreover, at first glance the response here enjoys some modal force. For many, pushing the guy just can't be permissible. This invites the hypothesis that the intuition is an application intuition, where semantic norms dispose one to apply the concept IMPERMISSIBLE to this case because the case presents conditions sufficient for the concept to apply.

I think evidence from disagreement and close inspection of the putative modal force of this intuition suggest that something else is going on. Consider what disagreement amounts to for the Gettier-type case. If I say Jones does not know, and another says that Jones does know, two options are salient: (i) we are employing concepts, governed by the same semantic-value fixing norms, and one of us has erred semantically; or (ii) we are employing concepts that are not governed by the same semantic norms, and so in this sense fail to disagree. Under option (ii), one is live to the possibility that the putative disagreement is merely verbal.

Substantive ethical intuitions do not fit this picture well. Consider a scenario where a consequentialist claims that pushing the fat man off the bridge is permissible. Rather than accept one of the two options above, those who have the opposing intuition of impermissibility are actually motivated to find a theoretical disagreement with the consequentialist. Both parties to a debate like this want to have a real disagreement with those who would guide action in an unacceptable way. Neither side of the debate wishes to charge the other with semantic error or to say the other side is simply talking about something else altogether.

Relatedly, upon careful inspection the modal force of ethical intuitions is not like the modal force of application intuitions. Whereas Jones cannot know that it is 12:30 PM, it might be permissible to push the fat man off the bridge. In saying this, I don’t mean to deny the following: if it is impermissible to push him off the bridge, then necessarily it is impermissible to push the fat man off the bridge. I am not talking about a modal force that can be inferred in this way. I am talking about the modal force built in to certain intuitions. This kind of modal force in the Gettier case intuition is different in kind than the modal force in the footbridge intuition. To be sure, the phenomenological difference between a semantic ‘must’ and a non-semantic compulsion is very subtle. It shows up more clearly when we consider cases of disagreement and the possibility of error, as above.

The emerging thought is that, for application intuitions, which show up outside of substantive ethics, semantics requires certain responses, whereas for ethical intuitions something else is compelling certain responses. In the ethical cases one might suggest that it is non-normative commitment that does the compelling, driving application of normative concepts just as semantic norms drive application in certain non-ethical cases.

If I am right to distinguish ethical intuitions from other philosophically interesting intuitions in this way, the result is significant not only because it would lead us to type ethical intuitions differently, but also because the semantic story that vindicates the epistemic properties for application intuitions would not apply to ethical intuitions. Indeed, we have a story for why it might seem as though ethical intuitions enjoy the same epistemic properties as intuitions of other sorts even though they do not. It is actually quite difficult to see how intuitions driven by antecedent, non-cognitive normative commitments could enjoy the kind of epistemic respect standard ethical theorizing accords them. If pressed, this might lead to a debunking account of the role ethical intuitions have played in theory building. At this point such a conclusion would be premature, however, for there
might be good reasons to rely on intuitions in theory construction that have nothing to
do with their justificatory statuses.

So far we have set forth self-evidence theory’s characterization of self-evident propositions in terms of their epistemic features, and examined two accounts of what self-evident propositions are, or what the corresponding intuitions would be, that try to explain these epistemic features. Note that the presumption here is that certain propositions, be they general and about abstract principles, or particular and about concrete cases, are self-evident. There is one other approach (not necessarily inconsistent with the above) I should mention before moving on.

2.3. INTUITIVE PERCEPTION

Some philosophers in the self-evidence tradition think there is something like intuitive perception of moral properties or facts when faced with particular, concrete moral situations. Audi, following Clarke, suggests that the properties apprehended are fittingness and unfittingness relations (Audi 2008: 482).

However the properties are characterized, those suspicious of sui generis non-natural properties would be right to express their concerns at this point. What are these moral properties, such as fittingness and unfittingness relations? And are we now to posit a sui generis faculty of Intuition by which we detect them? Such properties and faculties are metaphysically suspicious.

One reply here is to claim that perception of moral properties makes use of general cognitive capacities, and to claim that the properties or facts detected might well be natural ones. This debate will ultimately turn on issues in moral metaphysics, and, sadly, those issues would take us well beyond the purview of this paper.7

Having touched some criticism of self-evidence theory, we could now introduce some recent, influential concerns that come from the empirical literature. But let me hold off on this until we have the other main approach to intuitional moral epistemology on the table.

3. Seeming State Theory

It is fair to say that seeming state theorists begin by characterizing intuitions non-epistemically: for A to have an intuition that p is for A to have a mental state whereby it seems to A that p. Michael Huemer endorses such a view. According to him, ethical intuitions are initial, intellectual seemings, which can be contrasted with other species of seemings, e.g., experiential seemings (it seems that I have two hands), memorial seemings (I seem to remember turning the oven off), and introspective seemings (it seems that I now have a desire for ice cream) (Huemer 2005: 102). What unites all these kinds of seemings, according to Huemer, is that they are all seeming-type attitudes taken toward propositional contents (Huemer 2007: 30; see also Tolhurst 1998).

The epistemological bit of the theory is this: if it seems to A that p, then A has some justification for believing that p (in virtue of the seeming state) unless or until A acquires a defeater for that justification. Huemer puts this in terms of Phenomenal Conservatism (PC): ‘If it seems to S that p, then, in the absence of defeaters, S thereby has at least some degree of justification for believing that p’ (Huemer 2006: 148). One interesting thing about such a view is that it might bootstrap the epistemic principle, rendering it justified in the very way the principle describes. Why? Well, note that an attempt to reject it because it seems false would be self-defeating.
3.1. INTELLECTUAL SEEMINGS

This initial characterization suggests that seemings are of a kind, and in virtue of being of that kind they all justify in the same way. Critics of the view take issue (Bedke 2008; DePaul 2009). On the one hand, consider experiential seemings, where what makes something seem to be the case is the content of one’s experiential state. When it seems to me that I have two hands, for instance, it is some presentational content as of having two hands that constitutes its seeming to me that I have two hands. The point of putting it this way is to highlight the fact that the content would have to change, as opposed to just the attitude I am taking toward that content, for it to no longer experientially seem to me that I have two hands. The seeming is in the content, as it were.

Contrast this with its seeming to me that Jones does not know that it is 12.30 PM. One can have the very same content before one’s mind, and yet it might not seem to one that Jones doesn’t know that it is 12.30 PM. So when it does seem to me that Jones doesn’t know it is 12.30 PM my intuition is unlike the experiential seeming of having two hands in that the Gettier seeming isn’t in the content under consideration. We can add to this that the Gettier intuition also features some modal force, and is likely the product of, if not constituted by, the tacit recognition that one’s semantic norms require one to categorize this as determinately not knowledge.

What about ethical intuitions? They are not content-based, for one can have ethical content before one’s mind that is typically intuitive, but fail to have any corresponding seeming state. And, as indicated in our discussion of self-evidence theory, one’s substantive ethical intuitions are probably not driven by semantic norms. In light of this, it has been suggested that ethical intuitions are seeming states wholly constituted by certain phenomenological features like felt veridicality of ethical contents (Bedke 2008). One important consequence is that beliefs justified on the basis of such states would be distinguishable from justification based on other kinds of seemings, so defeaters or skepticism regarding ethical intuitions need not spread to all seeming-state-based justification.

3.2. PHENOMENAL CONSERVATISM

Some argue that seemings do not justify, or at least the seemings in ethical intuitions do not justify. This critique has comes from externalist quarters, who take note that seeming states are internal affairs.

Consider the following putative counterexample to Phenomenal Conservatism, offered by Peter Markie:

Suppose that we are prospecting for gold. You have learned to identify a gold nugget on sight but I have no such knowledge. As the water washes out of my pan, we both look at a pebble, which is in fact a gold nugget. My desire to discover gold makes it seem to me as if the pebble is gold; your learned identification skills make it seem that way to you (Markie 2005: 356–57).

For Markie, the source of the seeming is relevant to the justification the seeming confers, so it is not the seeming alone that secures defeasible justification.

At this point, an epistemic internalist is likely to note that it is not at all clear that you lack as much justification for your belief as I do for mine if our two seeming states are indistinguishable in their internal features (where features are internal if and only if they strongly supervene on the mental), if you have no idea of the suspect source of your seeming state, and if you have no other doubts about your seeming state. Having clarified
the case in this way, why wouldn’t you believe the pebble to be gold? There might be other categories of epistemic appraisal that differ across these two cases, those concerning warrant, knowledge, etc., but the seeming state theorist might still have something important to say about justification, or a kind of it.

4. Ethical Intuitions Empirical Literature

The empirical literature on intuitions is relatively new and it has received a lot of attention lately. Most of it is critical of intuitional epistemologies. Let me relay a bit of the research here. The discussion will be a bit hodgepodge, but I will organize it with headings that indicate the type of research done (roughly) and the kind of conclusions researchers have drawn (roughly).

4.1. Behavioral Studies and Intuitions as Emotional Responses

Some studies show that moral judgments, or at least some class of moral judgments, are influenced by emotions. The influence of disgust is particularly well studied. In Wheatley and Haidt (2005), for instance, subjects were hypnotized to feel disgust when reading words like ‘take’ and ‘often’. Given a description of an act or character with these words, hypnotized subjects registered more severe moral evaluations than non-hypnotized subjects. Even more surprising, some hypnotized subjects evaluated a perfectly permissible action as somewhat morally wrong. In another set of studies by Schnall et al. (2008), a dirty environment tended to elicit more severe moral judgments than a clean environment, at least for subjects that were highly aware of their own body states.

In light of experiments like these and dumbfounding effects, i.e., the inability to justify an initial moral judgment, Johnathan Haidt sees support for a hypothesis about what intuitions are: a great many intuitions are affective responses that result in moral judgments, and deliberative reasoning typically comes after that fact to provide post hoc rationalizations of those affect-backed judgments (Haidt 2001; Haidt and Bjorklund 2008). Clearly, such a view of intuitions will have consequences for their epistemic role (or lack thereof).

4.2. FMRI Studies and Intuitions as Unreliable

Other studies have shown emotional influence on moral judgment. Joshua Greene, for instance, uses fMRI technology to look at what areas of the brain are most active when subjects report moral judgments (Greene et al. 2001; Greene 2008). He finds that regions of the brain associated with more emotional, “intuitive” processing are more active in subjects who report characteristically deontological judgments—for example, the intuition that it is impermissible to push the fat man off the bridge in the footbridge version of the trolley case—whereas regions of the brain associated with deliberative cognition are more active in subjects who report consequentialism-friendly judgments. Moreover, Greene reports that cognitive load, which is the burden of being taxed with tasks that interfere with deliberative cognition, selectively interferes with consequentialism-friendly responses. So subjects who give the consequentialism-friendly response to Footbridge while under cognitive load show longer response times (Greene et al. 2008).

Both Joshua Greene and Peter Singer have used results like these to argue that moral intuitions, or at least deontology-friendly intuitions, are likely unreliable (Singer 2005; Greene 2008). Why unreliable? First, deontological intuitions are sensitive to morally irrelevant factors, such as whether you kill someone by pushing him versus throwing a switch.
Second, deontological intuitions are the result of fast and frugal “intuitive processes” that were selected by evolution to produce fitness-enhancing behaviors in some environment, rather than to track the moral facts (which perhaps this helps to explain the first thread of reasoning) (see also Timmons 2008).

What are we to make of these results and the conclusions drawn by the researchers? First, it is not clear whether intuitions as characterized by Audi, Shafer-Landau, Huemer and others are elicited in these studies, though this can be obfuscated by the fact that psychologists call the fast and frugal processes responsible for some moral judgments “intuitive processes”.

We should try to make sure that the propositions or mental states to which intuitional epistemologists attribute epistemic roles are the ones under study before we draw any negative conclusions about the justificatory role of intuitions from these empirical results.

Supposing that these studies do elicit intuitions as intuitional epistemologists have understood them, still, any claim like Greene’s and Singer’s that deontology-friendly intuitions are sensitive to morally irrelevant factors, or are the product of something selected by evolution without regard to the facts, stakes some claim on what is and what is not morally relevant, or what the moral facts are. These commitments are not the result of empirical inquiry, but intuition, or so it seems. So even if some intuitive justification is defeated when one appreciates these findings it is in part because intuition succeeds in justifying beliefs about what is or is not morally relevant, or beliefs about the moral facts. Also, the status of being a defeater is something the empirical results do not deliver – at best they exploit the fact that we treat certain considerations as defeaters (for similar points, see Berker 2009).

4.3. SURVEY STUDIES AND INTUITIONS AS UNRELIABLE

Other studies have shown that certain ethical intuitions are unreliable due to framing effects (as discussed in Sinnott-Armstrong 2008). Let me mention two such effects, one based on how questions are described, and another based on the order in which questions are presented.

In Petrinovich and O’Neill (1996) subjects were given trolley-type cases, where one must choose between the greater number and the lesser number. If given an option couched in terms of ‘kill’ wording (e.g., ‘to throw the switch which will result in the death of the one innocent person’), subjects on average disagree slightly with the action, whereas if given an option in terms of ‘save’ wording (e.g., ‘to throw the switch which will result in the five innocent people on the main track being saved’), subjects on average slightly agree with the action (Petrinovich and O’Neill 1996).

That same paper reported other studies that asked subjects to indicate their approval of actions (described in terms of saving) in three trolley-type questions given in sequence. One set of subjects got the questions in one sequence, and another set of subjects got the same questions in reverse sequence. On average, subjects more strongly approved of an action when it appeared first in the sequence than when it appeared last, and the degree of approval for the middle case varied depending on which of the other two trolley-type cases came first (Petrinovich and O’Neill 1996).

From framing effects like these, Sinnott-Armstrong argues that it is reasonable to ‘assign a large probability of error to moral intuitions in general’ (Sinnott-Armstrong 2008: 99). For him, any given moral intuition must be shown to be immune from such effects before it can be justified. (His case for this conclusion will be amplified in the next section).
The question about eliciting moral intuitions as proponents of intuitional epistemology understand them looms large here, too, for the studies above ask for approval of action rather than classifications of actions under moral categories. It is relatively easy to see why one would be more likely to approve of an action described in positive language (‘save’) than when described in negative language (‘kill’). People tend to approve of positive things and disapprove of negative things, and people are generally sensitive to other people’s evaluations (word choice helps to communicate how other people are evaluating things). Without similar explanations for explicit moral categorization, one wonders whether there would be framing effects.

Let us grant for the sake of argument that some framing effects on intuitions can be shown. Can we conclude from this, perhaps supplemented with some sense that the framing phenomenon will generalize, that moral intuitions generally are unreliable, and so any given intuition is probably incorrect? That is too quick. One basic problem concerns the move from a few cases where moral intuitions have been shown to be unreliable to the conclusion that moral intuitions are generally unreliable (as a class). This is a key step in Sinnott-Armstrong’s argument, for it allows him to infer for any given intuition that it is not to be trusted unless some confirming arguments are brought to bear.

The problem is twofold. First, the small sample of cases considered might have special characteristics that prevent us from projecting a general framing effect. In particular, it is relatively easy to see why word choice might have a pronounced influence on responses to dilemmas. Trolley-type cases and doing-allowing distinctions like those in the above studies involve already difficult decisions, and word choice might incline us one way rather than another. But how would word choice affect intuitions on something like a Rossian prima facie duty, e.g., that an action would fulfill one of my promises is a moral reason to do it (all else equal), or Sidgwick’s axiom of the Universal Good?

Second, it is noteworthy that Petrinovich and O’Neill (1996) report another set of trolley-type cases that did not show order effects (discussed in Sinnott-Armstrong 2008: 61). And even where order effects were shown, the effect wasn’t to change people from approval of action to disapproval of action, but to shift the degree of approval for action. So it might be misleading to talk of unreliability in this context. In the end, it isn’t clear how strong and pervasive the effects are, particularly for philosophers who are relying on intuitions for theory construction. Combine this with the first aspect of the problem, which might cloister whatever effects we have shown, and we lack good evidence that ethical intuitions as a class are unreliable due to framing effects.

4.4. OTHER CONSIDERATIONS AND INTUITIONS AS UNRELIABLE/IN NEED OF INFERENTIAL CONFIRMA-
TION

Actually, Sinnott-Armstrong thinks that there is other evidence showing that ethical intuitions as a class are likely unreliable. In his 2006 he argues that the following considerations should lead one to mistrust intuitive judgments without some non-intuitive corroboration of their contents: (i) which moral view one adopts affects one’s self-interests, so one is likely going to be biased when adopting a moral view; (ii) there is often moral disagreement without reason to think one party is more likely to be right than the other; (iii) emotions influence one’s moral view, and emotions cloud judgment; (iv) moral views are likely to be formed in circumstances conducive to illusion, and shown by heuristics and framing effects; and (v) moral beliefs might be the product of morally suspicious sources, e.g., if they are the result of the influences of the rich and powerful (Sinnott-Armstrong 2006). His point is that, at least cumulatively, these influences, or potential
influences, require us to gather some confirming evidence for the contents of our intuitions before those intuitions are justified (or before they justify beliefs). This does not necessarily sound a skeptical note. For Sinnott-Armstrong, the take-home message is that ethical intuitions are at best non-inferentially justifying.

Michael Huemer is impressed by similar considerations: that ethical intuitions conflict, that they have been influenced by one’s culture, that they have been influenced by evolutionary pressures, that they sometimes support theories that promote one’s self-interests, that they sometimes line up with strong emotion, and that the abstract ones are prone to an overgeneralization bias (Huemer 2008). On these grounds he mistrusts substantive ethical intuitions quite generally as likely unreliable, including, for example, in the intuition that infanticide is impermissible (Id., 382). Formal intuitions, like the transitivity of value (example 6), are given pride of place in his intuitional epistemology (Id., 386).

Rather than tackle these difficult issues in any detail, let me take a step back and note a problem I see for any reliability-based critique, viz., the lack of some independent, non-intuitional epistemology against which we can check the deliverances of intuition. In ethics what we start with and what we have to go on are our intuitions. Though we may recognize that not everyone agrees with what we find intuitive, and that our intuitions have been subject to certain forces, like all our other mental states, this doesn’t change the fact that each one of us has some evidence that certain moral propositions are true and others false. All we can do is consult the evidence we have, which comes in the form of intuition (cf. Ross 1930: 40–1; Brandt 1979: Chapter 1). Given this, how is any given person to discover that his/her substantive intuitions as a whole are unreliable (as opposed to some subclass)?

That said, I think there is something to reliability-based critiques as applied to historically influential Ethical Intuitionism. Recall that the historical position combined, inter alia, an intuitional epistemology and a non-natural metaphysics (as Huemer does). If we do this it turns out that moral intuitions as a class are not likely to be reliable. For, presumably, intuitions are part of the natural order, and caused by the natural order, whereas their objects are being conceived of as non-natural facts. What are the chances that nature causes intuitions in us that line up with the non-natural ethical facts of the matter? An alignment between the two would be a cosmic coincidence, one that we are not justified in believing. Given these low chances of alignment, any intuitive justification for ethical beliefs in non-natural properties or facts is defeated. So intuition might be more comfortably paired with a naturalist metaphysics (see Bedke 2009).

5. Conclusion

No doubt a lot has been left uncovered by this survey of issues, but I hope to have communicated some of the main lines of inquiry in the contemporary debate. Let me close by plugging an area that is fairly underdeveloped: the role intuition might play in expressivist theories, that is, theories of moral thought and language according to which moral judgments are not beliefs about the external world. Some of what I’ve said here pushes us in this direction. For instance, the claim that the modal force of ethical intuitions is not driven by semantic norms, but rather by normative commitments; and the claim that intuitional epistemology is uncomfortably paired with non-naturalist metaphysics. Work has begun in this area (see, e.g., Gibbard 2008), but like so much of what I have discussed, it is ripe for further exploration.
Acknowledgements

Thanks to Mark Timmons, Mark van Roojen, Oisin Deery, and an anonymous referee from Philosophy Compass for comments on a previous draft.

Short Biography

Matthew S. Bedke is an Assistant Professor of Philosophy at the University of British Columbia. His work focuses on the nature of normativity, both practical and epistemic. His articles include ‘The Iffiest Oughts’ (Ethics 119, 2009), ‘Ethical Intuitions: What They Are, What They Are Not, and How They Justify’ (American Philosophical Quarterly 45, 2008) and ‘Intuitive Non-Naturalism Meets Cosmic Coincidence’ (Pacific Philosophical Quarterly 90, 2009). Bedke received a JD is law from the University of Arizona in 2001, and a PhD in philosophy from the University of Arizona in 2007.

Notes

* Correspondence: Department of Philosophy, University of British Columbia, 1866 Main Mall E370, Vancouver, BC V6T 1Z1, Canada. Email: mbedke@interchange.ubc.ca

1 For more on the relation between this base-level justification and considerations of coherence, see DePaul (2005) and Huen (2008). For a discussion of intuition’s role in moral theory that is not explicitly epistemic, see Copp (forthcoming).

2 Other Rossian prima facie duties include the duty of reparation, the duty of gratitude, the duty of justice, the duty of beneficence, the duty of self-improvement, and the duty of non-maleficence (1930: 21).

3 It is not clear to me that Audi would endorse this view. He might think that containment is a non-semantic, non-conceptual relation between two properties. If so, it is harder to see how possession of the concepts that refer to these properties would allow one to see their containment relations via mere reflection, and so harder to see why self-evident propositions would play a cognitive role that differs from non-self-evident propositions. Audi will likely appeal to some version of intuitive perception, which I reference below.

4 Thanks to commenters from Philosophy Compass for pressing this view.

5 By contrast, if presented with a hypothetical case of aqueous stuff, stuff that is clear, potable, flows around objects, has greater volume in its solid state than in its liquid state, etc., I would be inclined to classify it as water. But this classification would not carry modal force. No matter how water-like, I wouldn’t think this stuff must be water except to express my incredulity at the thought this it is not water. In fact, there is information I could be given that would incline me to believe that it just can’t be water, viz., that it is not H2O. My dispositions to apply the concept WATER in the first case would not be driven by semantic norms governing WATER, but it would in second case, or so I think.

6 Others have endorsed a view like this (Kaupinnen 2007: 101; Sosa 2007: 101; see also Goldman 2007).

7 For recent discussions of moral perception, not necessarily intuitive, see Audi (2010), Dancy (2010), Väyrynen (2008), and McBrayer (2010).

8 Because some emotional areas of the brain are activated for subjects who give consequentialism-friendly responses, Greene now says that the deontological intuitions are more alarm-like, whereas the consequentialist intuitions are more currency-like (Greene 2008: 41).

9 Haidt and Bjorklund define a moral intuition as ‘the sudden appearance in consciousness or at the fringe of consciousness, of an evaluative feeling (like-dislike, good-bad)...’ (2008: 188).

10 Rather than say intuitions are unreliable, I should say that beliefs based on ethical intuitions (perhaps in some circumstance-type C) are disproportionately false.

11 Sinnott-Armstrong defines an intuition as ‘a strong immediate moral belief’ (2008: 47), which does not map on well to the other characterizations we have considered.

12 Sinnott-Armstrong says that, if Ross is right, then intuitions on principles are arrived at via intuitions on particular cases, and these have been shown to be vulnerable to framing effects (Sinnott-Armstrong 2008: 73). But, again, not all concrete intuitions have been shown to be vulnerable to framing effects that make them unreliable.
Works Cited


