

Moral Intuition in Philosophy and Psychology

To appear in Neil Levy and Jens Clausen (eds.), *The Springer Handbook of Neuroethics*.

Antti Kauppinen

Trinity College Dublin

Draft, December 19, 2012

Almost everyone agrees that ethical disputes cannot be resolved by appealing to empirical evidence alone. How, then, can we settle which moral principles to adopt or what to think of a particular moral issue? It is not uncommon for philosophers to appeal to *intuitions* as fundamental evidence or source of evidence. Is this practice justifiable, particularly in the light of recent findings in psychology and neuroscience?

To begin with, we have to know what we are talking about. The term ‘intuition’ has many related uses in philosophy and psychology. Generally speaking, intuition contrasts with reasoning, and has connotations with spontaneity and insight. But closer examination reveals that there are a number of different phenomena in this area, with the result that we must take great care to avoid talking past one another. My goal in this chapter is to clarify these distinctions and relate different conceptions in different disciplines to each other, in the hope of understanding better when and why we may rely on intuitions in ethical inquiry.

The most obvious distinction is between intuition as a kind of psychological *faculty* and intuition as a psychological *state*. In the first sense, intuition contrasts with reason and vision, for example. It is something we can make use of when forming beliefs. Some Early Modern Rationalists may have believed that we have such a special capacity to gain immediate access to abstract truths. But few if any contemporary philosophers believe in any such faculty, so this sense can safely be left aside here.

The crucial questions thus concern the nature and significance of *intuition as a psychological state*.¹ What is it for someone to have a moral intuition that it is wrong to kill one in order to save five, for example, and what does it matter if they do? For philosophers, I claim, intuitions are primarily identified by their *putative epistemic and dialectical role*. In particular, intuitions are supposed to play a *foundational* or *quasi-foundational* role in non-empirical justification. They are justifiers that require no further proof, or at least initially credible starting points of inquiry. As such, they aren't theory-based, and constitute data theories have to account for or explain away.

So the question is: what if any psychological states are fit to play such a justificatory role and why? Philosophers differ sharply. For *Intuitionists*, intuitions are either intellectual appearances, attractions to assent to propositions, or beliefs that result from merely adequately understanding their content. For *Rawlsian coherentists*, intuitions are considered judgments that are inputs to a process of reflective equilibrium. For *intuition sceptics*, in contrast, there is nothing that plays the intuition-role – so-called intuitions are just beliefs or inclinations to believe that have no special justificatory status (Williamson 2007, Cappelen 2012).

Psychologists and other empirical scientists, in turn, at least typically think of intuitions as beliefs that result from automatic, non-conscious and non-rational psychological processing. I will label such psychological processes *intuitive processes*, and call the beliefs that result from them *intuitions** to distinguish them from intuitions in the sense that philosophers use. It turns out to be very important to understanding and addressing philosophical claims made by psychologists and neuroscientists that they are not always talking about the same thing as philosophers when the latter use the word 'intuition'. Yet at

¹ Sometimes the term 'intuition' is also used for the *proposition* that is the *content* of the psychological state. I will leave aside this use here. The adverb 'intuitively' and the adjective 'intuitive' are used much more broadly, and shouldn't be taken to entail that anyone has an intuition (*pace* Cappelen 2012).

the same time, psychological results may nevertheless be very important for understanding intuitions in the philosophers' sense.

I will begin with an overview of intuitions as understood in empirical moral psychology, and briefly examine some claims made on the basis of empirical findings. I'll then outline the various philosophical conceptions, and in the final section try to reconcile the two pictures somewhat.

1. Intuitions in Empirical Moral Psychology

As noted, psychologists typically define intuitions as beliefs we acquire without (conscious) reasoning (that is, intuitions*). For example, according to Jonathan Haidt, a moral intuition can be defined as

the sudden appearance in consciousness of a moral judgment, including an affective valence (good-bad, like-dislike), without any conscious awareness of having gone through steps of search, weighing evidence, or inferring a conclusion. (Haidt 2001, 818)

It is an empirical question which of our moral judgments are intuitions* in this sense, and which judgments result from reasoning, for example. The primary goal of empirical moral psychology here is to find out to what extent intuitive processes *explain* why people judge as they do, and what role they play in moral agency. But both psychologists themselves and experimental philosophers have also drawn normative or epistemic conclusions from these explanations.

1.1 Dual Process Theory

It has proven useful in psychology to make a distinction between two kinds of psychological process. Some processes are fast, effortless, non-conscious, uncontrolled and non-intentional,

often associative or pattern-matching and non-linear, and frequently affective. They are also often contained within relatively self-contained mental ‘modules’ such as visual perception that are innate and typically products of natural selection. These processes form what is often called *System 1* or the Intuitive System. In contrast, the processes that belong to *System 2* or the Reasoning System are relatively slow, effortful, conscious and attention-demanding, controlled, intentional, serial, and potentially rational. The evidence in favour of this contrast between conscious and non-conscious thinking is by now overwhelming (see Wilson 2005 and Kahneman 2011).

From a philosophical perspective, it is crucial that System 1 comprises of very *heterogeneous* processes. Some System 1 processes are likely to result in false beliefs, especially when triggered outside the context in which they are adaptive. The heuristics and biases approach (Kahneman, Slovic, and Tversky 1982) has extensively studied such predictably irrational processes. However, other System 1 processes are likely to produce true beliefs. For example, I understand ordinary English sentence in a flash, as a result of non-conscious processing I can’t control. Yet I, of course, *know* what it means. So sometimes the outputs of System 1 are likely to be false, and sometimes they are likely to be true. It all depends on which System 1 processes are at issue, and what the circumstances are like. Bearing this point in mind is very important in thinking about the epistemic relevance of intuitive processes.

1.2 Intuitions and Explanation

Why do we make the moral judgments we do? One kind of answer to this explanatory question concerns the *proximal* causes of moral judgments – the psychological states and processes that cause moral judgments. They are responses to the agent’s (perceived) situation. A further explanatory question concerns the origins of the proximate processes: why do

certain situational features trigger certain responses? The answers to this question are *distal* explanations.

It is a signature claim of much recent empirical moral psychology that the proximal causes of many, indeed most, judgments are automatic and often affectively laden intuitive processes, so that many of our moral beliefs are intuitions* (see Pizarro, this volume). Briefly, we're often unable to articulate our reasons, emotion-related areas of the brain are activated when we judge, and we evaluate quickly, constantly, and without taxing working memory. As a rule, System 2 becomes involved only when there's a problem – when intuitions conflict or the social context requires explicit articulation (Haidt 2001). No one denies that *some* moral judgments result from reasoning – though reasoning often serves to rationalize pre-existing intuitions* after the fact.

So intuitive processes appear to proximally explain many of our moral judgments. What is the *distal* explanation for the intuitions* that we have? Many contemporary psychologists appeal to the *evolutionary benefits* of intuitions*. For example, according to Haidt's Moral Foundations Theory (Haidt 2012), natural selection has favoured the development of six innate modules that produce the various affective responses underlying intuitions*. Roughly, the story is that groups whose members respond negatively to things like causing innocents to suffer and disrespecting those with high social status, and positively to helping those in need or respecting the elders are likely to outcompete groups whose members do not have such responses. Consequently, the genes that program for the adaptive responses become more prevalent by group selection. Different cultures then fine-tune these innate responses in different ways.

For my purposes, the details of this broadly plausible account do not matter. The important questions arise from the assumptions that intuitions* are hard-wired and evolutionary fitness-enhancing.

1.3 Are Intuitions Unreliable?*

In the psychological literature it is typical to assume that moral beliefs that result from explicit reasoning are epistemically unproblematic. To be sure, from a philosophical perspective this is highly contingent – after all, there is such a thing as bad reasoning and false premises! But what should we think of the epistemic status of intuitions*?

One popular claim is that intuitions* are unreliable, because they result from *affective* processes. The best-known such argument is made by Joshua Greene (e.g. Greene 2008). According to Greene, only *nonconsequentialist* beliefs result from intuitive processes, while consequentialist beliefs are the result of reasoning. On the basis of this, he constructs an a posteriori argument for consequentialism (cf. Kauppinen (forthcoming a)):

1. Empirical investigation shows that nonconsequentialist moral intuitions* are proximately caused by intuitive emotional reactions.
2. Intuitive emotional reactions don't justify the beliefs they cause, because they are sensitive to morally irrelevant features.
3. So, empirical investigation undermines the justification of nonconsequentialist moral intuitions*.
4. Nonconsequentialist moral theory rests crucially on nonconsequentialist intuitions*.
5. So, nonconsequentialist moral theory is epistemically unsupported.

Greene's critics, such as Berker (2009), have pointed out that the second premise is not an empirical or neuroscientific one. But as Greene rightly counters, empirical evidence in favour of the first premise still does crucial work in the argument. Specifically, he believes that it shows nonconsequentialist intuitions* result from negative emotional responses towards using personal force to harm others (Greene et al. 2009). The distal explanation for this is roughly that such reactions facilitated peaceful coexistence in small groups in human prehistory. The

evolved ‘point-and-shoot morality’, however, is likely to misfire in modern conditions, so we shouldn’t rely on intuitions* in moral thinking (cf. Singer 2005). Greene’s openly normative assumption is that whether we use personal force in causing harm is morally irrelevant. As he notes, this is something it is hard to for anyone to deny.

There are, however, at least three major problems with this line of argument. First, there are nonconsequentialists who disavow appeal to intuitions about cases, most notably Kantians (e.g. Wood 2011). Even if most people’s nonconsequentialist beliefs were based on unreliable gut reactions, it doesn’t follow that nonconsequentialist theory isn’t true or unsupported, as long as there is some other kind of justification for it. Second, even if consequentialist beliefs result from reasoning, it doesn’t follow that they are justified, unless the premises of that reasoning are themselves justified. And at least some of those premises appear to rely on intuitions, such as the intuition that it is better to save more rather than fewer people, at least other things being equal. Nonconsequentialists are free to turn the tables and say that this intuition is a part of an evolved point-and-shoot morality that sometimes gets it right and sometimes doesn’t. If, on the other hand, being the product of evolution shows that a moral belief is unjustified – which many now consider a red herring (see e.g. Kahane 2011) – then the consequentialist is hoist by his own petard.

Finally, many have called into question the specific interpretations of empirical data that Greene offers (e.g. C. Klein 2011). Even if intuitions* are results of an affective process, it has certainly not been shown that all morally relevant emotions are responsive to features that are uncontroversially morally irrelevant. For example, the sort of reactions that classical sentimentalists thought central to morality, such as resentment and gratitude, are responsive to features such as being used as a mere means for another’s ends or exceeding expectations. Nothing so far shows they couldn’t confer justification to beliefs (see below).

1.4 Can Intuitions* Be Reliable?

As I noted earlier, there are System 1 processes that are potentially sources of knowledge. If (some) intuitions* result from some such process, they will be trustworthy. I'll discuss two recent proposals to this effect.

According to what I'll call the *Expert Intuition View*, moral intuitions* *can* be reliable in just the same way as what are called expert intuitions are, provided we have the right kind of training. Allman and Woodward (2008) argue that moral intuitions* are the output of a species of affective social cognition that can be trained to be responsive to moral features by the same sort of implicit learning that teaches nurses to recognize which infants are sick, for example. In this kind of learning, the learner is exposed to cues she may not be able to articulate, forms judgments, and receives independent feedback that tells her whether her judgments are on the right track or not – for example, a child's temperature returns to normal in response to treatment (G. Klein 1998). With enough experience, her System 1 delivers functional 'intuitions' about what to do.

The chief, and in my view fatal, problem with the Expert Intuition View is that one of the necessary conditions for implicit learning, namely immediate and unambiguous feedback (see Kahneman and Klein 2009), is missing in the case of morality. If I act on the mistaken intuition* that it's acceptable for me to break a promise to a student in order to get a bit of rest, what is the equivalent of an infant's fever getting worse? Nothing. Even if moral mistakes reliably have bad consequences for *others*, no one thinks there is reliably a negative signal for the *agent*. People can persist with mistaken moral views quite easily, especially if surrounded by the like-minded.

A different approach to trustworthy intuitions* is provided by the Moral Grammar View (Dwyer 1999, Mikhail 2011; Hauser 2006). According to it, the process that yields

moral intuitions* is non-conscious and automatic, but strictly rule-governed and computational rather than associative and affective. This innate moral competence is analogous to Chomsky's universal grammar. Again like linguistic understanding, the rules that govern this System 1 process are inaccessible to ordinary users, but can be articulated by experts, who deduce the existence of rules like the Doctrine of the Double Effect on the basis of observational data about judgments.

This is no place to evaluate Mikhail and Hauser's explanatory theory. Supposing it to be true, what can we conclude about the epistemic status of intuitions*? Mikhail talks freely about knowledge and competence, and seems to assume that beliefs that result from exercising competence are somehow correct. But there's a crucial disanalogy to language here: Moral judgments purport to represent how things are, and are not in any sense 'correct' just because they are entailed by *some* system of rules. An error theorist about ethics might happily endorse the Moral Grammar View as showing that we are hard-wired to make systematically false moral judgments. So the epistemic status of the Doctrine of the Double Effect, for example, remains a mystery, if this explanation is correct.

2. Intuitions in Moral Philosophy

Although there is some controversy about how common appeals to intuition are in philosophy in general (Cappelen 2012), there is little doubt that they play a major role in contemporary normative ethics. Intuitions (or intuited propositions) about either particular cases or general principles are treated as presumptively valid starting points of ethical inquiry. As W.D. Ross put it, "the moral convictions of thoughtful and well-educated people are the data of ethics just as sense-perceptions are the data of a natural science" (Ross 1930/2002, 41). Normative theories typically aim to capture intuitions, or else find some way of undermining the

authority of a particular intuition. But views about the nature and importance of intuitions differ widely.

2.1 *Self-Evidence Intuitionism*

The classical Intuitionist view is that certain moral truths are *self-evident*, much in the way that mathematical axioms are to those who understand them. (Sometimes the term ‘Intuitionism’ is also used for a related metaphysical thesis that moral properties are non-natural, but I’ll restrict myself to the epistemological use.) Here’s Ross:

That an act *qua* fulfilling a promise, or *qua* effecting a just distribution of good . . . is *prima facie* right, is self-evident; not in the sense that it is evident from the beginning of our lives, or as soon as we attend to the proposition for the first time, but in the sense that when we have reached sufficient mental maturity and have given sufficient attention to the proposition it is evident without any need of proof, or of evidence beyond itself. It is evident just as a mathematical axiom . . . is evident. (Ross 1930, 29–30)

On a natural reading, for me to have the intuition that fulfilling a promise is *prima facie* right is for me to find the proposition evident (believe it to be necessarily true?) simply on the basis of understanding it properly and attending to it. According to Robert Audi’s contemporary reformulation, an intuition is a cognitive state whose content one doesn’t infer from what one believes or any theory but which one forms on the basis of an adequate understanding of the intuited proposition (Audi 2004, 33–36). Its content is a *self-evident* proposition. A proposition is self-evident “provided 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” (ibid., 49).

Audi thus considers intuitions as *beliefs* that are individuated by their distinctive *justification and aetiology*. Ernest Sosa’s (2007) related rationalist view in general epistemology differs in that he considers an intuition to be an *attraction to assent* to a proposition rather than a belief. This is because we can have an intuition that *p* even if we

know that p is false – for example, all the lemmas of a paradox are intuitive. On either picture, given that mere understanding suffices for knowing their content, intuitions are instances or sources of *a priori* knowledge. Given that their content is not tautological, they are sources of knowledge about *synthetic* truths.

Non-moral examples of putatively self-evident propositions include *Nothing is both red and blue all over*, *No vixens are male*, and *The existence of great-grandchildren requires at least four generations of people*. Moral Intuitionists differ on what kind of moral propositions are self-evident. For Sidgwick, only the most *fundamental moral principles* can lay claim to self-evidence. According to him, they are 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”, and that “as a rational being I am bound to aim at good generally” (Sidgwick 1907, 382). From these he infers the truth of a form of Utilitarianism.

Ross, in contrast, believes that several *mid-level moral principles* are self-evident. For example, he believes it is self-evident that we have (overridable) moral reason to keep our word, match rewards to desert, be grateful to benefactors, to refrain from injuring others, and to benefit others. These are among considerations that always weigh in favour or against performing an action. An agent’s overall or final duty in a particular case is a function of these *pro tanto* reasons. Finally, while most Intuitionists agree that *verdicts about particular cases* cannot be self-evident, H. A. Prichard and moral particularists deny this. We can just see what the right thing to do in a particular situation is. Indeed, Prichard thought this is epistemically prior: “If I cannot see that I ought to pay this debt, I shall not be able to see that I ought to pay a debt.” (Prichard 2002, 4)

Recent work by Audi, Shafer-Landau (2003), and others has dispelled many standard objections to the Self-Evidence View. There is no appeal to a special faculty of intuition, only

ordinary understanding. A proposition can be self-evident even if some people who understand it don't regard it as self-evident, certain, or even true – after all, all that matters is that they *would be justified* in believing it on the basis of mere understanding. Nor need a self-evident proposition be obvious, as adequate understanding may take time and effort. Given the preceding, it is not surprising if people *disagree* about self-evident propositions – even if they adequately understand them, they may be led to deny them as a result of indoctrination, bad theory, or self-interest. A person suffering from such issues might sincerely believe that she has an intuition – but insofar as her belief or attraction isn't justifiable by mere understanding of the content, she is mistaken about its nature.

Nevertheless, challenges remain. Although epistemological Intuitionism is logically independent of non-naturalist moral metaphysics, they are *de facto* allies. After all, if moral facts were natural facts, why couldn't we come to know them the way we come to know other natural facts? But if moral facts are non-natural and thus causally inefficacious, while intuitions are psychological states with causal histories governed by natural laws, it would have to be a fantastic cosmic coincidence for the contents of the intuitions to align with the non-natural facts (Bedke 2009). Further, mere adequate understanding of the content is supposed to justify belief in the self-evident proposition. Sometimes this makes sense. What it is to understand the concept of a vixen is, at least in part, to know that it does not apply to males. So it is no surprise that merely understanding it suffices for knowing the *conceptual* or *analytic* truth that no vixens are male. But how can mere understanding reach to *synthetic* truths that are not about relations between concepts? That turns out to be very hard to account for (for recent attempts, see Jenkins 2008 and Bengson 2010).

2.2 Seeming-State Intuitionism

In general epistemology, it has recently become popular to think of intuitions as intellectual appearances or seemings (Bealer 2000, Chudnoff 2011). This view has adherents in moral epistemology as well, as Michael Huemer's definition of moral intuition shows:

An intuition that *p* is a state of its seeming to one that *p* that is not dependent on inference from other beliefs and that results from thinking about *p*, as opposed to perceiving, remembering, or introspecting. (Huemer 2005, 102)

Huemer's definition has two parts: an intuition is a) a seeming and b) the result of merely thinking about the proposition. Seemings or appearances in general are *non-doxastic*, propositionally contentful states: it may seem to me that the stick in the water is bent, even though I do not *believe* that it is. They have a *presentational phenomenology*: when we have them, it's as if we're directly presented with their objects or truth-makers. Consequently, they are *compelling*: they attract assent to their content and appear to rationalize belief. Some such seemings are perceptual, such as my visual experience of having a computer in front of me. But others are, according to seeming-state intuitionists, *intellectual*: they result from merely thinking about the proposition. They claim that merely thinking about killing one in order to save five can give rise to a non-doxastic, presentational, and compelling experience of moral wrongness.

Suppose that when I merely think about it, cheating on my spouse seems morally wrong to me. Does this justify my believing so? Seeming-state intuitionists tend to subscribe to a view about justification called epistemic liberalism (Bengson 2010). According to this view, we are justified in taking things to be as they appear to be, unless we have sufficient reason to doubt the appearances. Not every experience is a seeming in the specified sense, so the view doesn't license belief in just anything we dream or fantasize about. But when it comes to genuine seemings, we're not epistemically blameworthy for taking them at face value, other things being equal.

Walter Sinnott-Armstrong (2006) argues against this kind of view in ethics. According to him, empirical evidence shows that other things are not equal: the way things seem to us, morally speaking, is often biased by partiality, beset by peer disagreement, clouded by emotion, subject to framing effects, and influenced by disreputable sources like religion. Such factors, in general, give us reason to doubt appearances of a particular kind, so we can't take them at face value without some sort of further confirmation. But the need for further confirmation, Sinnott-Armstrong says, means that intuitions are unfit to play a foundational role in justification, so Intuitionism is false.

2.3 Coherentism and Intuitions

What are often described as intuitions also play a crucial role in the best-known coherentist approach to moral epistemology, the method of *reflective equilibrium* (Rawls 1971, Daniels 1980). Here is how Rawls describes it:

People have *considered judgments* at all levels of generality, from those about particular situations and institutions up through broad standards and first principles to formal and abstract conditions on moral conceptions. One tries to see how people would fit their various *convictions* into one coherent scheme, each considered judgment whatever its level having a certain *initial credibility*. By dropping and revising some, by reformulating and expanding others, one supposes that a systematic organization can be found. (Rawls 1974, 8, my emphasis).

As I have highlighted, Rawls talks about considered judgments and convictions as the initially credible (but potentially revisable or dispensable) starting points of moral inquiry. In this lightweight sense, intuitions need not be any special sort of mental state or have any particular kind of aetiology. The emphasis on *considered judgments* rules out unreflective gut reactions, however.

While the deflationary aspect of Rawlsian intuitions has its attractions, it also raises an immediate epistemic question. Why should the mere fact that we believe something yield *any* initial credibility to the believed proposition? Precisely because the aetiology of the beliefs

doesn't enter the picture, considered ideological or self-serving judgments seem to start out with the same status as rational insights. Coherentists might respond by waging that such beliefs fall out in the process, but insofar as it is path-dependent (that is, the outcome depends on what the inputs are), there is no guarantee that the outputs of reflective equilibrium aren't systematically biased.

Coherentists might appeal to the notion of *wide* reflective equilibrium, in which psychological, sociological, and other empirical facts are brought into the balancing act. This might rule out beliefs with some intuitively problematic aetiologies, such as beliefs based on knee-jerk reactions or sensitive to the use of personal force (see the discussion of Greene above). But why? Because these causal histories typically result in beliefs that do not fit with the rest of our moral beliefs. This means that if our moral beliefs are distorted to begin with, widening the reflective equilibrium won't help. The ideologically brainwashed will regard the process we regard as brainwashing as conducive to true beliefs.

Perhaps the best response to worries about both Seeming-State Intuitionism and coherentism is provided by Mark van Roojen (forthcoming). He acknowledges that the kind of considerations Sinnott-Armstrong puts forward may mean that intuitions don't suffice to justify belief on their own. But when a proposition is *both* the content of a moral intuition *and* coheres with other intuitive propositions, belief in it will be justified. The initial credibility provided by intellectual appearance is, as it were, confirmed by coherence. Since reflective equilibrium is applied selectively only to appearance-based beliefs, low quality inputs to the balancing process are filtered out at least to some degree.

3. Reconciling the Pictures?

How do the psychological and philosophical views of moral intuitions relate to each other?

Could both disciplines learn something from each other? Before trying to answer these

questions, let me map out the different views of intuition that I have discussed:

Views About Moral Intuition	<i>Type of mental state</i>	<i>Aetiology</i>	<i>Claimed epistemic standing</i>
Psychology	Belief	System 1	Most consider dubious
Self-Evidence Intuitionism	Belief or attraction to assent	Mere adequate understanding of content, which also justifies belief	Constitutes or is a source of knowledge of non-natural facts
Seeming-State Intuitionism	Appearance/ seeming/presentation	Merely thinking about the content	Justifies belief in the absence of reason to doubt
Coherentism	Belief/conviction	Reflection, consideration	Initial credibility

Table 1. Views About Moral Intuition

As Table 1 makes obvious, many intuitions* are not intuitions in the philosophical sense. This has several consequences for contemporary debates in empirically informed ethics or experimental philosophy.

First, as I pointed out, not all System 1 processes are epistemically equal. One kind of intuitive process that is particularly likely to issue in false beliefs is quick gut reaction. But for philosophers, it is emphatically not the case that the truth of an intuited proposition supposed to be manifest in a quick flash. It is *immediate* only in the sense that its justification is not mediated by some further, itself justified belief. Coming to an adequate understanding of a proposition or thinking about a thought experiment is an effortful, System 2 process that may take time, often compared to what it takes to appreciate a work of art (e.g. McMahan 2013).

As Audi (2004) puts it, an intuition can be a conclusion of *reflection*, although it can't be a conclusion of *inference*. This kind of reflection is not the sole privilege of philosophers, but it is one of the things they're trained to do. This gives a positive reason to regard philosophers' intuitions as superior.

Second, it is often claimed that surveys or questionnaires reveal people's moral intuitions about particular cases, so that by varying the cases and performing statistical analyses we can discover what ordinary people's intuitions are sensitive to (Kahane 2012). But as John Bengson (2012) has pointed out, this is simply not a valid inference. Even if responses reflect intuitions*, it doesn't follow that the subjects have an intuition in any sense that interests philosophers. This at least limits the usefulness of survey studies and complicates any empirical study of intuitions.

Intuitive Sentimentalism

Although not all intuitions* are intuitions in the philosophical sense, the latter are nevertheless the outcome of some intuitive process rather than reasoning or inference. Could empirical evidence about the non-conscious functioning of brain and mind help understand intuitions and their role? I believe it is going to be one important ingredient in the mix. I will finish by briefly sketching the case for thinking of intuitions as *manifestations of moral sentiments* (see Kauppinen (forthcoming b)).

Start with the observation that moral intuitions appear to have a *distinctive and diverse phenomenology*. When something *seems* wrong, it often *feels* wrong – even if one *believes* it is right. This suggests that intuitions are non-doxastic experiences, as Seeming-State Intuitionists think, but not the same kind of experience as other intuitions. Second, moral seemings can directly *motivate* us to act and react. This is clearest in cases like Huck Finn's lying to slave-catchers: although he didn't *believe* that it was the right thing to do (it went

against everything he was taught), it nevertheless *felt* like the right thing, and this was sufficient to move him to act. Third, note that apparently manipulating the subjects' emotions results in change in their intuitions (e.g. Valdesolo and di Stefano 2005), while it is unlikely to change the mathematical or other run-of-the-mill intellectual intuitions.

All these special features of moral intuitions are readily accounted for if they are emotions. Which emotions? Those that manifest moral sentiments. Sentiments, I take it, are dispositions to feel, act, think, and believe (cf. Prinz 2007). The sentiment of liking one's dog manifests itself in delight on seeing the animal, sadness when it is ill, noticing when its favourite food is for sale, desire to buy the food, and so on. The sentiment of moral disapprobation towards cheating on one's spouse manifests itself in anger towards someone who cheats on his spouse, guilt when one thinks of doing it oneself, desire to refrain from acting on certain desires, and so on. In this context, I've argued, the emotional manifestations of the sentiment constitute moral appearances: the anger you feel towards the unfaithful husband *presents* his action as morally wrong and attracts you to *believe* that it is morally wrong, while also having a distinctive phenomenal feel and motivational role. Since these emotions constitute moral seemings, they confer defeasible initial credibility to their contents just as other seemings do.

But aren't emotions subject to general epistemic defeaters, due to their characteristic fickleness and partiality? My response is that not all emotions are created equal. *Canonical moral sentiments* are felt from what Hume called 'The Common Point of View' – roughly speaking, they result from a process of impartially sympathizing with the hedonic states and reactive attitudes of those affected by the actual or hypothetical action. When my anger or guilt is based on this kind of sentiment, the generic reasons for doubting emotional appearances are absent – the emotional reactions aren't fickle, rash, partial, or ill-informed. Further, such responses are not sensitive to what everyone agrees are morally irrelevant

features, such as mere physical distance, but rather to features like being treated as mere means or not receiving equal reward for equal contribution, which are plausibly morally relevant.

So let's go back to the questions we started out with. I have the moral intuition that it is wrong for a doctor to grab a random person off the street and take his vital organs to save five others. According to the view I've just sketched, this intuition is trustworthy when it consists in an emotional response that I have when I merely think about the case in the canonical way – not just understanding the proposition, but also imagining myself in the shoes of those affected by the action. This suggestion is compatible with the well-supported empirical hypothesis that emotions play a crucial causal role in moral judging. Although the response itself results from a System 1 process, it may be preceded by conscious, System 2 effort to reflect on the situation and empathize with different perspectives. Given that the sentimental intuition presents the action as wrong and isn't subject to standard defeaters, it is fit to play at least a quasi-foundational role in moral justification. So it seems to me that if we consider at least some moral intuitions to consist in a sharply delimited kind of emotional response, we can go fairly far in reconciling the psychological and philosophical conceptions without sceptical consequences.

References

- Allman, J. & Woodward, J. (2008). What Are Moral Intuitions and Why Should We Care About Them? A Neurobiological Perspective. *Philosophical Issues*, 18, 164-185.
- Bealer, G. (2000). A Theory of the A Priori. *Pacific Philosophical Quarterly*, 81, 1-30.
- Bedke, M. (2009). Intuitive Non-Naturalism Meets Cosmic Coincidence. *Pacific Philosophical Quarterly*, 90, 188-209.
- Bengson, J. (2010), *The Intellectual Given*. Dissertation, University of Texas at Austin.

- Bengson, J. (forthcoming). Experimental Attacks on Intuitions and Answers. *Philosophy and Phenomenological Research*.
- Berker, S. (2009). The Normative Insignificance of Neuroscience. *Philosophy and Public Affairs*, 37 (4), 293-329.
- Cappelen, H. (2012) *Philosophy Without Intuitions*. Oxford: Oxford University Press.
- Chudnoff, E. (2011), What Intuitions Are Like. *Philosophy and Phenomenological Research*, 82 (3), 625-654.
- Daniels, N. (1979). Wide Reflective Equilibrium and Theory Acceptance in Ethics. *Journal of Philosophy*, 76 (5), 256-282.
- Dwyer, S. (1999). Moral Competence. In K. Murasugi & R. Stainton (eds.), *Philosophy and Linguistics* (pp. 169-190). Boulder, CO: Westview Press.
- Greene, J. (2008). The Secret Joke of Kant's Soul. In W. Sinnott-Armstrong (ed.) *Moral Psychology. Vol. 3*. (pp. 35-80). Cambridge, MA: MIT Press.
- Greene, J.D., Cushman, F.A., Stewart, L.E., Lowenberg, K., Nystrom, L.E., & Cohen, J.D. (2009). Pushing moral buttons: The interaction between personal force and intention in moral judgment. *Cognition*, 111 (3), 364-371.
- Haidt, J. (2001). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review*, 108, 814-834.
- Haidt, J. (2012). *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. New York: Pantheon.
- Hauser, M. (2006), *Moral Minds*. New York: Harper Collins.
- Huemer, M. (2005). *Ethical Intuitionism*. New York: Palgrave MacMillan.
- Jenkins, C. (2008). *Grounding Concepts: An Empirical Basis for Arithmetical Knowledge*. Oxford: Oxford University Press.
- Kahane, G. (2011). Evolutionary Debunking Arguments. *Noûs*, 45 (1), 103-125.
- Kahane, G. (forthcoming). The Armchair and the Trolley. *Philosophical Studies*.
- Kahneman, D. (2011). *Thinking Fast and Slow*. New York: MacMillan.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kauppinen, A. (forthcoming a). Ethics and Empirical Psychology. In M. Christen, J. Fischer, M. Huppenbauer, C. Tanner & C. van Schaik (eds.) *Empirically Informed Ethics*. Springer.
- Kauppinen, A. (forthcoming b). Intuition and Belief in Moral Motivation. In Gunnar Björnsson et al. (eds.) *Moral Motivation: Evidence and Relevance*. Oxford: Oxford University Press.
- Klein, G. 1998. *Sources of Power. How People Make Decisions*. Cambridge, MA: MIT Press.

- Klein, C. (2011). The Dual Track Theory of Moral Decision-Making: a Critique of the Neuroimaging Evidence. In *Neuroethics*, 4, 143-162.
- Mikhail, J. (2011), *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. Cambridge: Cambridge University Press.
- Prichard, H. A. (2002). *Moral Writings*. Ed. Jim MacAdam. Oxford: Oxford University Press.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rawls, J. (1974). The Independence of Moral Theory. *Proceedings and Addresses of the American Philosophical Association*, 48, 5-22.
- van Roojen, Mark (forthcoming), Moral Intuitionism, Experiments, and Skeptical Arguments. In A. Booth and D. Rowbottom (eds.) *Intuitions*. Oxford: Oxford University Press.
- Ross, W. D. (1930/2002). *The Right and the Good*. Ed. Philip Stratton-Lake. Oxford: Oxford University Press.
- Shafer-Landau, Russ (2003). *Moral Realism: A Defence*. New York: Oxford University Press.
- Sidgwick, H. (1907). *Methods of Ethics*. 7th Edition. London: MacMillan & Co.
- Singer, P. (2005), Ethics and Intuitions. *The Journal of Ethics*, 9, 331-352.
- Sinnott-Armstrong, W. (2006). Moral Intuitionism Meets Empirical Psychology. In Terry Horgan & Mark Timmons (eds.), *Metaethics After Moore*. Oxford: Oxford University Press.
- Sosa, E. (2007), *A Virtue Epistemology I: Apt Belief and Reflective Knowledge*. New York: Oxford University Press.
- Valdesolo, P. and DeSteno, D. (2006), Manipulations of Emotional Context Shape Moral Judgment. *Psychological Science*, 17 (6), 476-477.
- Williamson, T. (2007). *The Philosophy of Philosophy*. Blackwell: Oxford.
- Wilson, T. D. (2002). *Strangers to Ourselves: Discovering the Adaptive Unconscious*. Cambridge, MA: Harvard University Press.
- Wood, A. (2011), Humanity as an End in Itself. In Derek Parfit, *On What Matters. Volume 2*. Oxford: Oxford University Press, 58-82.