Intuitions, emotions and gut reactions in decisions about risks: towards a different interpretation of ‘neuroethics’

Sabine Roeser

To cite this article: Sabine Roeser (2010) Intuitions, emotions and gut reactions in decisions about risks: towards a different interpretation of ‘neuroethics’, Journal of Risk Research, 13:2, 175-190, DOI: 10.1080/13669870903126275

To link to this article: http://dx.doi.org/10.1080/13669870903126275

Published online: 12 Feb 2010.
Intuitions, emotions and gut reactions in decisions about risks: towards a different interpretation of ‘neuroethics’

Sabine Roeser*

Philosophy Department, Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands

(Received 30 September 2008; final version received 4 February 2009)

Empirical research by Jonathan Haidt and Joshua Greene seems to support the idea that in moral decision-making under uncertainty, people follow their initial intuitions and ‘gut feelings’. Rational judgements are at most rationalizations or afterthoughts in our judgements about risks. This paper will challenge the theoretical assumptions made by Greene and Haidt, by proposing a different theory of ethical intuitions and emotions. Ethical intuitions and emotions should not be conflated with spontaneous ‘gut reactions’. Rather, ethical intuitions and emotions can be the source and the result of ethical reflection and deliberation. This allows for different interpretations of the empirical findings of Haidt and Greene and of psychologists who study emotional responses to risks, such as Paul Slovic and George Loewenstein. Emotional and intuitive responses to risk should not be seen as heuristics that are prone to be biases; rather, they should be seen as invaluable sources of insight when it comes to judging the moral acceptability of risks.

Keywords: risk; ethics; emotions; intuitions; neuroethics; dual process theory

1. Introduction

In recent years, research on the role of affect, feeling and intuitions in decision-making under uncertainty has become very influential. This work seems to be supported by empirical research by Jonathan Haidt and the neuroscientist Joshua Greene. All these approaches endorse a dual process theory according to which reason and emotion are distinct faculties that have opposite tasks. Intuitions and emotions are considered to be gut reactions that work like heuristics, but that are highly unreliable. In this paper I offer a different interpretation of the empirical findings by Greene and Haidt by challenging the dichotomy between reason and emotion that they presuppose. I present an alternative theoretical framework according to which ethical intuitions and emotions should be distinguished from gut reactions. Intuitions and emotions can be the source and the result of ethical reflection and deliberation. This makes room for a different view on the role of intuitions and emotions in decisions about risks. I will argue that when judging the moral acceptability of risks, ethical intuitions and emotions are an invaluable source of insight.

*Email: S.Roeser@tudelft.nl
2. The role of affect in decision-making under uncertainty

In the last years, there has been an increasing interest in the role of affect in decision-making under uncertainty. Paul Slovic and his colleagues have developed a theory about the so-called ‘affect-heuristic’: affective responses to a large degree determine our judgements about risk. If we have a positive attitude towards a hazard, people judge its benefits as high and its risk as low, and vice versa (Finucane et al. 2000; Slovic et al. 2002, 2004). According to Slovic and his colleagues, this affect-heuristic can mislead us, but on the other hand, affect is important in order to grasp the meaning of events that numbers often fail to convey (Loewenstein et al. 2001 make similar claims). Cass Sunstein, though, is much more dismissive of affect. He considers it as a major source of flaws in our thinking about risks and uncertainty (Sunstein 2005).

Most authors in this field assume that emotions (or affect) and rationality are distinct sources of insight that have opposite tasks. These authors endorse the theoretical framework of dual process theory (DPT). According to DPT, there are two distinct systems through which human beings apprehend reality: System 1 is emotional, affective, intuitive, spontaneous and evolutionary prior; System 2 is rational, analytical, reflective and occurred later in our evolution. System 1 processes function as ‘heuristics’ that help us navigate relatively smoothly through a complex world without requiring too much time, but they are also ‘biases’, because they might let us overlook important complexities and nuances. System 2 processes are more reliable, but they come at the cost of being slower since they require more reflection (cf. Epstein 1994; Sloman 1996, 2002; Stanovich and West 2002). In a similar vein, the neuropsychologist and moral philosopher Joshua Greene has argued, based on functional magnetic resonance imaging (fMRI) scans, that different regions in the brain are activated in assessments of different sorts of moral dilemmas. ‘Emotional regions’ are invoked in situations that involve personal moral judgements (e.g. hurting somebody who is close by) and ‘cognitive regions’ are invoked in situations that involve impersonal moral judgements (e.g. hurting somebody who is far away; cf. Greene et al. 2001, 2004). Similar ideas are developed by Jonathan Haidt (cf. also Greene and Haidt 2002). In a groundbreaking article, Haidt (2001) has developed a so-called ‘social-intuitionist model’ (SIM). His claim is that our moral judgements are formed by spontaneous, intuitive gut reactions, and our rationality at most works as a rationalization post-hoc, being the ‘rational tail’ that is wagged by the ‘emotional dog’. Affect is what primarily steers us. Haidt goes even further than defenders of DPT in denying rationality an important normative role. The framework of DPT seems to be ambiguous as to the status of affect in judgements about risks. It can allow for an important role of affect (this is emphasized by Slovic), but it can also serve to claim that we should invoke reason as a final arbiter of our affectively formed judgements (Sunstein invokes DPT in order to be dismissive of emotions). Haidt’s approach, though, makes it doubtful whether this is possible at all. According to Haidt, reasoning resembles a lawyer more than a judge: reasoning means arguing for an emotionally pre-established view rather than finding a true answer. However, it is unclear whether these ideas do sufficient justice to certain emotions, i.e. moral emotions. Such emotions can be a source of reflection and practical rationality. They are not necessarily spontaneous but they can also be intertwined with complex, long-lasting narratives (cf. Roeser 2009). In addition, doubts have been raised whether Haidt’s approach covers all forms of moral judgement (cf. Fine 2006; Pizarro and Bloom 2003; I will discuss this further on). I have criticized DPT in another paper (Roeser 2009). In this paper, I will take a closer
look at Haidt’s and Greene’s work and will argue that their findings can be interpreted in a different way by understanding intuitions and emotions not as contrary to, but as a specific form of rationality. This alternative interpretation can be applied to emotional judgements about risks, which will allow emotions to play a more important normative role in judgements about risks than the one that they are assigned to by Slovic, Loewenstein and Sunstein.

3. A neurological basis for our ethical intuitions and emotions

Based on fMRI scans, Greene has shown how different brain regions are involved in different kinds of ethical decision-making. Greene has asked subjects to judge various ethical dilemmas such as trolley cases. The majority of research subjects judged an overt moral wrongdoing to be condemnable, but they found it acceptable or preferable to allow something wrong to happen. Greene used various trolley problems. Trolley problems have become very popular in the philosophical and psychological literature on moral judgements. They were initially conceived of by Philippa Foot (1967) and further developed by Judith Jarvis Thomson (1976, 1985). Trolley problems are always variations on the following theme: a trolley speeds down a runway. You see that further down the track, there are five people who will get killed if the trolley hits them. The only way to avoid this is for you to: a push a button which diverts the trolley to a different track that has one person on it who will be killed, but the five on the initial track will be saved; b push a large man down the track who is standing next to you on a footbridge. Where most people tend to find a morally acceptable, they find b morally unacceptable. In impersonal cases, we tend to be utilitarians (maximize aggregate goodness); in personal cases, we tend to be Kantians (respect for persons: don’t use another person merely as a means to achieve a good outcome). There are many other cases that cover the terrain between these two possibilities a and b. These alternative cases might help to give us a better understanding of what the crucial differences between our intuitions are (cf. Hauser 2006, 110–20; Kamm 2007). I will come back to this further on. In any case, Greene’s fMRI studies have shown that in the impersonal cases (like a), cognitive regions in our brain are invoked. In the personal cases (like b), with people who judged them unacceptable (the vast majority), emotional regions are invoked (Greene and Haidt 2002; Greene et al. 2001, 2004). The few subjects who did find b-like cases acceptable needed significantly longer to make their decisions, and cognitive regions were invoked instead of emotional regions. Greene concludes that utilitarian decisions are made by reason and deontological decisions by emotion. According to Greene, this can be explained by an evolutionary perspective: in our evolutionary past, we mainly had to make personal moral decisions, and these were made by our early, emotional brain regions. The further we evolved, the more complex moral decisions we had to make since our actions were able to have more and more far-reaching consequences. These decisions are made by our more evolved, rational regions in the brain.

Greene contends:

There is a growing consensus that moral judgments are based largely on intuition – ‘gut feelings’ about what is right or wrong in particular cases. (Greene 2003, 847)

This seems indeed to be the accepted view in contemporary moral psychology. However, I will show that moral philosophers who emphasize the importance of
intuitions have a different understanding of that notion than Greene, Haidt and other psychologists. In addition, I will look at cognitive theories of emotions. These alternative approaches offer different understandings of intuitions and emotions, which will allow for a different interpretation of the empirical findings by Greene and Haidt. This will accordingly allow for a different view concerning the role of emotions in risk perception than the views offered by Slovic, Loewenstein and Sunstein.

4. An alternative approach to intuitions and emotions
Jonathan Haidt (2001) invokes ethical intuitionists in order to explain his SIM. Haidt uses the notions intuition, gut reaction and emotion interchangeably. Haidt’s SIM is a form of Humeanism. However, intuitionism and Humeanism are two opposed meta-ethical theories. For ethical intuitionists, intuitions are not the same as ‘gut feelings’. Gut feelings are instinctive, spontaneous responses. Moral intuitions instead are understood by intuitionists as non-inferential moral judgements. Most intuitionists do not see intuitions as affectively loaded; rather, they think that moral intuitions are a product of our rational faculty. Furthermore, contrary to gut feelings, intuitions do not necessarily pop up spontaneously. They might be the result of a long process of reflection, but they are not inferentially or deductively based on the preliminary ingredients of reflection (cf. Ewing 1929; Prichard 1912). We need to know the relevant factual information about a situation in order to assess its moral value, but that assessment is not a deductive inference, since one cannot derive an ‘ought’ from an ‘is’ (the naturalistic fallacy; cf. Moore 1903/1988). This moral assessment can best be understood as an ‘intuition’ (cf. Roeser 2005a). In addition, the category of emotions is broader than that of gut feelings. The fact that a moral judgement is based on emotions does not exclude the possibility that it is also based on reflection.

Hence, it is important to maintain the philosophical distinctions between Humeanism and ethical intuitionism. Although there are some overlapping ideas between Haidt’s SIM and ethical intuitionism and cognitive theories of emotions, there are crucial respects in which these approaches differ, as I will show in the following subsections. And these differences can shed important new light on our understanding of intuitions and emotions in response to risks.

4.1. Intuitions
Ethical intuitionists think that all moral judgements are based on ethical intuitions (or basic moral beliefs). Intuitions are non-inferential judgements. They are not the result of a deductive argument. However, ethical intuitions should not be conflated with gut reactions. Many scholars take ‘gut reactions’ to be unconscious, irrational and unreflected (cf. Haidt 2001). In contrast, ethical intuitionists think that intuitions can be reflective and justified, even if we are not able to provide for further going justification in the sense of a deductive argument.

William Alston (1989) distinguishes between being justified and having justification, and between first- and second-order beliefs. He and other externalists in epistemology defend that being justified in holding first-order beliefs does not require having second-order, justifying beliefs. This corresponds well with the ideas of ethical intuitionists. They think that some moral beliefs are basic, they cannot be justified any further without invoking an infinite regress, circular reasoning or arbitrary assumptions (also cf. Alston 1993). In addition, Thomas Reid contends that even if it is in principle
possible to provide further arguments for some moral beliefs, a person can nevertheless be justified in holding these beliefs without being able to provide for such an argument (Reid 1788/1969). Most intuitionists are commonsense philosophers. Their framework makes it possible to take our everyday moral judgements as genuine knowledge.

Of course this gives rise to various questions: are ethical intuitions infallible? How to deal with disagreement? Can we revise our intuitions?

Intuitions are not infallible. This is a common misunderstanding amongst opponents of intuitionism. All intuitionists have emphasized that although we cannot avoid grounding our moral views on basic moral beliefs or intuitions, those are no guarantee to truth. They invoke the analogy with mathematical knowledge and sense perception. In both cases, we have to take something as a starting point, i.e. either axioms or sense data. But our supposed understanding of axioms can be flawed, and our sense data too, due to malfunctioning of our faculties or unfavourable conditions. According to Reid, this is even more the case with our moral faculty; he thinks that it can be easily manipulated and corrupted (Reid 1788/1969, 371). Nevertheless, intuitions are the starting point of all ethical reflection (Reid 1788/1969, 234). This still allows for the possibility that intuitions can be revised, for example, in the light of contrary evidence, such as in a disagreement. However, even then we will have to invoke intuitions, albeit different ones than before. This can be the case if we undergo a formative experience that changes our basic beliefs. In that case, we have not used a deductive argument; rather, a moral proposition appears in a different light due to a different understanding of a situation.

That we can revise our moral intuitions becomes even more interesting in the light of some empirical findings that are discussed in an article by Cordelia Fine (2006; a similar argument is made by Pizarro and Bloom 2003). In a response to the work by Jonathan Haidt, Fine discusses empirical research that shows that people invoke second-order judgements with which they keep their unfavourable initial judgements, such as racist or sexist prejudices, under control. According to Fine, this shows, in contrast to Haidt’s claims, that rational judgements do not necessarily follow gut feelings but that they can actually supersede them. In line with the common understanding of the relation between spontaneous and reflected judgements, Fine interprets the initial judgements as intuitive gut reactions and the second-order judgements as rational, non-intuitive. However, based on the framework of ethical intuitionism, the second-order judgements can still be understood as intuitions. They can be basic moral beliefs in the sense of being non-inferential. I think that one of the best examples of a basic moral belief is ‘all human beings have equal worth’ . It seems impossible to find a non-question-begging argument for this. We can point to the nasty implications of racism and sexism, but these implications are only nasty in the light of the idea of equal worth. Alternatively, one can think of arguments such as that all human beings are equally talented, nice, smart, etc., but these are empirically false. In addition, they don’t seem to be able to provide us with the reason why everybody has equal worth. The principle of equal worth can serve as a second-order judgement with which people keep their (e.g. sexist or racist) gut reactions in check, and it can best be understood as a basic moral belief or intuition, in the sense of being non-inferential. Hence, Fine’s examples might be counterexamples to Haidt’s SIM, but not to ethical intuitionism.

4.2. Emotions

Ethical intuitionists are rationalists. They think that moral discourse is objective, and only rationality can track objective moral truths. To make matters even more complex,
let me introduce yet another understanding of ethical intuitions. My own view is as follows: ethical intuitions are paradigmatically cognitive moral emotions. This involves a combination of ethical intuitionism with a cognitive theory of emotions. I call this approach ‘affectual intuitionism’ (Roeser 2002, 2006b, 2010). To use the example of the principle that all human beings have equal worth: an emotion such as sympathy can help us understand this principle. By feeling with and for another person, I understand that she has the same needs and rights as I do. Sympathy can broaden our perspective; we share a ‘circle of concern’ with another person, as Martha Nussbaum (2001, 319) calls it.

Greene and Haidt endorse a Humean view about the relation between affect and reason in moral judgements. Their own metaethical outlook is also Humean. The Humean view about the relation between reason and emotion is still predominant with metaethicists, also by those who reject Hume’s metaethical views. For example, the special issue of Philosophical Explorations that is devoted to the discussion of the meaning of neurological findings for moral philosophy is set up within the dichotomy ‘sentimentalism’ versus ‘rationalism’ (cf. Gerrans and Kennett 2006). Either moral experience is emotional, but then it must be subjective (sentimentalism), or it is objective, but then it must be rational (rationalism). However, this strict dichotomy between reason and emotion is seriously questioned by many philosophers and psychologists who study emotions. Emotions are a form of cognition and insight, especially when it comes to evaluative judgements (Frijda 1987; Nussbaum 2001; Roberts 2003; Scherer 1984; Solomon 1993; Stocker and Hegemann 1996). Emotions allow us to be practically rational (Damasio 1994).

The cognitive theory of emotions that I endorse says that emotions have cognitive and affective aspects at the same time (cf. Roberts 2003; Scherer 1984; Zagzebski 2003). For example, prototypical emotions such as contempt, indignation, shame, guilt and demerit include a moral judgement (the cognitive aspect), but they also involve affective states, such as feeling ‘pangs’ of guilt. With such emotions, we cannot separate the cognitive from the affective aspect. They are two sides of the same coin. In the same vein, it is futile to ask whether the affective or the cognitive response comes first. Experiencing indignation means having a judgement and a feeling. Experiencing indignation is not the response to an initial, purely cognitive moral judgement. Forming the judgement and having the feeling go hand in hand. Based on factual information of a situation, we form a moral judgement that is cognitive and affective at the same time. This understanding of moral emotions is very similar to the understanding of moral judgements of ethical intuitionists. The difference is that they do not acknowledge the affective aspect of moral judgements. Intuitionists (as other rationalists such as Kantians) at most assign emotions a motivating force. However, affective states are not merely contingent add-ons that serve no epistemological value. Caring about something or being affectively engaged opens our eyes to morally salient aspects of situations (cf. Blum 1994; Little 1995).

To wrap up, this is how the various approaches discussed so far differ in their understanding of moral judgements, gut reactions, intuitions and emotions.

On the view as endorsed by Greene and Haidt, spontaneous moral judgements are irrational, unreflected gut reactions. Based on ethical intuitionism, we could say that this is not necessarily the case. Some spontaneous moral judgements might indeed be gut reactions in the sense of prejudices and stereotypes that should crumble on genuine reflection. However, other spontaneous moral judgements might be moral
intuitions that can sustain further going reflection, even though they are not based on such reflection, at least not in every case.

According to Cordelia Fine, second-order moral judgements are different from gut reactions, intuitions and emotions. In contrast, ethical intuitionism provides for the possibility that second-order judgements are intuitions. Basic moral beliefs can help overcome our own prejudices, stereotypical beliefs and gut reactions. This holds, for example, for the principle of equal worth. We cannot avoid invoking basic moral beliefs. The difference is that such a principle can sustain critical reflection, whereas stereotypes can’t.

My own version of ethical intuitionism allows that second-order moral judgements can be intuitions and emotions at the same time. Some of our moral intuitions and emotions might be like gut reactions, but not all are. These categories overlap, but they do not coincide. Hence, the fact that a moral insight is based on, or involves, intuitions and emotions does not necessarily mean that it is a gut reaction, and accordingly, that it is irrational.

This also means that the common framework to understand intuitions and emotions about risks does not suffice: intuitions and emotions transcend the categories of DPT. Intuitions and emotions help us to be practically rational; they are sources of knowledge. Hence, they have features of System 1 (e.g. being affective and non-inferential) and also of System 2 (e.g. being reflective and sources of knowledge and justification; for a detailed discussion of this in the context of risk perception, cf. Roeser 2009).

5. Neuro-intuitions revisited

In the previous section, I have sketched a different understanding of intuitions and emotions. This approach allows for a different interpretation of the empirical data from Greene and Haidt. Instead of a sceptical, naturalist understanding of ethical intuitions and emotions, I propose in this section to understand them as (fallible) sources of insights into objective moral truths. In Section 6, I will apply these ideas to intuitions and emotions concerning risks.

5.1. Neuro-intuitions and moral objectivity

Greene and Haidt argue that since moral judgements are nothing more than intuitive gut reactions, they are not so much attempts to capture moral truths but rather reflections of evolutionarily and socially formed prejudices. They might be beneficial to our survival but reveal nothing about normativity. According to Greene:

Understanding how we make moral judgments might help us to determine whether our judgments are perceptions of external truths or projections of internal attitudes. (Greene 2003, 849)

According to Greene, the fact that our moral judgements have a neural substrate indicates that they are not perceptions of external, objective truths:

But maybe this pair of moral intuitions has nothing to do with ‘some good reason’ and everything to do with the way our brains happen to be built. […] To make a long story short, we found that judgments in response to ‘personal’ moral dilemmas, compared with ‘impersonal’ ones, involved greater activity in brain areas that are associated with
emotion and social cognition. [...] Over the last four decades, it has become clear that natural selection can favour altruistic instincts under the right conditions, and many believe that this is how human altruism came to be. (Greene 2003, 848)

This leads Greene to formulate the following hypothesis:

We ignore the plight of the world’s poorest people not because we implicitly appreciate the nuanced structure of moral obligation, but because, the way our brains are wired up, needy people who are “up close and personal” push our emotional buttons, whereas those who are out of sight languish out of mind. (Greene 2003, 849)

However, I don’t agree. Being able to detect a neural substrate of a certain judgement does not entail that the judgement is no more than a projection of that neural substrate. Even if a judgement is a perception of an external truth, it has to be accompanied by a neural substrate. That a certain belief is evolutionarily hardwired does not mean that it cannot be true; to the contrary, it might be an adjustment to an insight that has proven to be appropriate.

Greene compares our moral intuitions with other effortless cognitive processes, for example, to detect whether a face is male or female, or whether somebody is sexy. Based on the latter analogy, Greene claims to have found the ‘beginnings of a debunking explanation of moral realism’ (Greene 2003, 849). Moral experience has a perceptual phenomenology, just like the experience of sexiness has. But, says Greene, the experience of sexiness can be explained by survival mechanisms to find suitable mates. Baboons find other baboons much sexier than, for example, Tom Cruise, whereas human beings will not be sexually appealed by baboons. There is no universal, objective truth about sexiness. Analogously, Greene concludes, we can explain the experience of morality by survival mechanisms without needing to invoke moral truths.

I wish to refrain from making any commitments as to the (non-)objectivity of sexiness judgements. For the sake of the argument, let us grant Greene that he is right that sexiness is merely an evolutionary construct with no objective truth to the matter. However, from this nothing follows as to the (non-)objectivity of ethics. Greene merely presents us with a possibility, but this is not a conclusive argument. His argumentative structure gets ad hoc when he considers the other example that he has invoked, namely that of gender perception. About this example, he says:

Note that according to this view moral judgment is importantly different from gender perception. Both involve efficient cognitive processes that give rise to a perceptual phenomenology, but in the case of gender perception the phenomenology is veridical: there really are mind-independent facts about who is male or female. (Greene 2003, 849)

This is of course begging the question, why should moral judgement not be compared with the objectivity of gender perception rather than with the supposedly evolutionarily construed perception of sexiness? Greene goes on to admit that more argument is needed, which I agree with. One possible objection he tackles is how he could defend increasing aid to the poor and reject moral realism at the same time. His ‘brief reply’ is:

Giving up on moral realism does not mean giving up on moral values. It is one thing to care about the plight of the poor, and another to think that one’s caring is objectively correct. (Greene 2003, 850)
This is philosophically unsatisfying. Why should we care about the poor? Surely, that is a question moral philosophers should try to answer. Greene helps himself to a moral intuition without giving a coherent justification of it, which is problematic, since in his view, moral intuitions aren’t justified. Greene first states that as a matter of fact, people care more for those close by than those far away, and that this can be explained by an evolutionary account (cf. the first quote from page 849). Furthermore, Greene claims that our moral judgements are nothing more than the product of evolutionary processes (cf. the quote from page 848). Greene owes us an account of how then he can ‘care about the plight of the poor’ (cf. the quote from page 850). He cannot claim that this is what our moral intuitions should be like (even though as a matter of fact they are not), as he has removed the possibility of moral critique from his account. Greene started his paper by saying that he does not think that we can derive a ‘moral ought’ from a ‘neural is’, and he is completely right in this view. However, in the remainder of the paper he does exactly this, thereby taking away the ground on which our moral gut feelings could ever be criticized, presumably by more elaborate moral intuitions. Hence, Greene’s approach faces serious problems concerning metaethical issues, and it is far from obvious that his findings preclude a realist interpretation of moral intuitions.

5.2. Neuro-intuitions and moral dilemmas

Let us now look at the more normative–ethical claims of Greene’s experiments with trolley problems. Greene, and based on Greene’s work, also Peter Singer (2005), argue that utilitarian considerations are superior and that people are irrational and inconsistent in their ‘respect for persons’ intuitions about personal dilemmas, and rational in their utilitarian judgements about impersonal dilemmas. However, that is question-begging towards long-lasting philosophical debates between utilitarians and deontologists. The conclusions Singer and Greene draw based on Greene’s studies about trolley problems are not as straightforward as they suggest. There are huge debates about the doctrine of double effect (cf. McIntyre 2004), the doing–allowing distinction (cf. Howard-Snyder 2007) and trolley problems (cf. Kamm 2007), and of course even more about utilitarianism versus deontology. I don’t have space in this paper to come even near to doing justice to the complexities of these debates. All I wish to show here is that Greene’s and Singer’s rejection of ‘respect for persons’ considerations is question-begging, to say the least. The fact that there are numerous articles by philosophers giving reasons and arguments for non-utilitarian answers to trolley-like problems refutes Greene’s and Singer’s presupposition that such answers are irrational, unreflected gut reactions. Even if one does not agree with the non-utilitarian answers, one has to admit that there is more to them than just a gut reaction that cannot sustain deliberation. The key philosophical question is of course: are these intuitions justified or are they incoherent and irrational? Greene and Singer reject our intuitions. Greene says explicitly, in line with Haidt’s views about rationality: ‘Deontology, then, is a kind of moral confabulation’ (Greene 2007a, 63).

Are Greene and Singer right in their rejection of our ethical intuitions? Look at the following example:

A surgeon walks into the hospital as a nurse rushes forward with the following case. ‘Doctor! An ambulance just pulled in with five people in critical condition. Two have a damaged kidney, one a crushed heart, one a collapsed lung, and one a completely
ruptured liver. We don’t have time to search for possible organ donors, but a healthy young man just walked in to donate blood and is sitting in the lobby. We can save all five patients if we take the needed organs from this young man. Of course he won’t survive, but we will save all five patients.’ (Hauser 2006, 32)

Marc Hauser contrasts the hospital case with the trolley case in which one can redirect the trolley to a track with one person and in doing so save five persons on another track (Case a). In the hospital case, the ‘respect for persons’ decision is the one that practically everybody finds preferable (don’t sacrifice one to save five), whereas in the specific trolley case with which it is compared, most people prefer the utilitarian decision (sacrifice one to save five). Note that the hospital case is similar to the footbridge case (Case b: you can save five people from a trolley by pushing a large man on the track who is standing next to you on a footbridge). However, I find the hospital case even clearer, maybe because the initial situation is more realistic (five dying patients versus one healthy person), which makes the possible solution (kill the healthy person to save the five patients) even more bizarre.

Hauser discusses three possible explanations for our diverging intuitions about trolley-like cases: the Kantian imperative that forbids using persons merely as means to an end, personal versus impersonal harm and the doctrine of double effect (Hauser 2006, 116–7). Hauser offers a taxonomy of trolley cases, refining them more and more to tease out the key differences. He concludes that the explanation of our diverging intuitions in various trolley-like cases is the doctrine of double effect:

It is impermissible to cause an intended harm if that harm is used as a means to a greater good. In contrast, it is permissible to cause harm if that harm is only a foreseen consequence of intending to cause a greater good. (Hauser 2006, 120)³

Note that this includes the Kantian idea of respect for persons. And in the examples that Hauser discusses, the ones that have a personal connotation are the ones that most people find unacceptable. So, it seems that in the end, all three considerations help understand what is going on in people’s moral intuitions and emotions when thinking about trolley cases.

Maybe Singer and Greene find the utilitarian solution to the hospital case acceptable, but I guess many people will disagree. Greene and Singer both argue that we should rise above our evolutionary determined intuitions and find a rational answer. But it is not obvious that the utilitarian answer is the rational answer in each and every case. Greene and Singer ignore the normative discussion in which intuitionists and utilitarians have been engaged for the last 300 years. To mention just a few objections of intuitionists against utilitarianism, it allows for exploitation of minorities, it ignores issues of fair distribution, it ignores the motives from which actions are done and it ignores agent relativism (I made a promise to you, Harry borrowed the book from Susan, he’s my son, etc.). In contrast with Greene and Haidt, intuitionists don’t make a descriptive claim about intuitions, but a normative claim: our ethical intuitions have to be taken seriously. The utilitarian Henry Sidgwick emphasized that the utilitarian principle is based on an intuitive insight, i.e. the principle of rational benevolence.⁶ Hence, also utilitarians cannot avoid ethical intuitions. Every ethical theory needs to invoke intuitions at some point (cf. Roeser 2005b).

However, most people have more diverse ethical intuitions than pure utilitarianism would allow. Trolley problems and similar dilemmas show that for most people there are cases where emotional–deontological concerns (supposing that Greene is right in
equating them – something Kantians won’t appreciate) trump ‘rational’–utilitarian concerns. Ethical intuitionism can make sense of this. Most intuitionists are pluralists.\footnote{Intuitionists have argued that there is a plurality of \textit{prima facie} principles that are instantiated in different ways in the world (Broad 1930/1951; Ewing 1929; Prichard 1912; Ross 1930/1967). Pluralism allows for the idea that in one situation Principle A is appropriate (e.g., utilitarianism), in another situation Principle B (for example, respect for persons). This is in direct contrast with the monism advocated by utilitarians and Kantians. According to them, there is one master principle that can be applied to every single case. Utilitarians and Kantians both have to reject our intuitive response to either of the two kinds of trolley-like cases (personal versus impersonal). Intuitionists don’t. They can point to the subtle difference between these two cases and say that this is what justifies the different responses. Hence, intuitionists can endorse our intuitions and the fine-grained analysis that Hauser offers, whereas Kantians and utilitarians have to reject one of the two conflicting intuitions.}

Since Greene sets up the normative–ethical spectrum as solely divided between consequentialists and deontologists, he fails to see a straightforward alternative theoretical explanation of our diverging intuitions, i.e. pluralism as developed by ethical intuitionists (Timmons 2007 raises the same point against Greene, explicitly referring to Ross and Prichard). Ethical intuitionism does not say that we should always follow our first impressions. Ethical reflection is also possible on ethical intuitionism. But this reflection also involves intuitions (cf. my discussion of Fine’s critique of Haidt). Greene says: ‘… it is exceedingly unlikely that there is any rationally coherent normative–ethical theory that can accommodate our moral intuitions’ (Greene 2007a, 72). It depends on what one means by rationally coherent. If Greene means by it a monist theory that can subsume all judgements under one principle, then intuitionists would agree with him. Most intuitionists are pluralists and think that we cannot find one overarching moral principle that covers all possible cases. However, we can give reasons for our diverging ethical intuitions in different cases. Maybe not everybody is equally skilled in articulating these differences, but people are nevertheless proficient in making these distinctions. Marc Hauser (2006) and John Mikhail (2007) draw the analogy with our capacity to use language. Many native speakers who are proficient in their language are not capable of explicating the rules of grammar that they are nevertheless perfectly capable of using. According to Hauser and Mikhail, the same might hold for our moral sense. Even though Hauser and Mikhail relate their work to John Rawls, in this respect they come as close to the intuitionist Reid as one could think. As said before, Reid thought that virtuous people are justified in their moral beliefs, even if they cannot provide further going arguments for them (Reid 1788/1969).

Moral philosophers might be the ones who are able to give a further going analysis of diverging moral intuitions. Indeed, in his comment on Greene, Mikhail offers very illuminating analyses of the trolley case versus the footbridge case (my Cases $a$ and $b$) that show that in the footbridge case, there are two stages where the agent would have to commit battery (pushing the man and letting the train hit him), as means to the end of saving the other people, whereas in the trolley case there is only one point at which one would commit battery (letting the train hit the man), and that is part of the side effects, not of the means to the end of saving the other people (Mikhail 2007, 86).

In addition, the examples that Greene has studied all involve dilemmas. Now it is a defining characteristic of dilemmas that there is not a clear-cut solution. Whatever horn of the dilemma we choose, we have failed in a sense by not going for the other
horn. A situation with one obvious best solution is by definition not a dilemma. Nevertheless, Singer and Greene treat the various dilemmas as if they do have a clear solution, thereby always unquestioningly favouring a utilitarian above a deontological solution. That is question-begging towards deontologists, to say the least. Hence, the fact that utilitarian decisions invoke rational parts of our brains, whereas deontological decisions invoke emotional parts of our brains does not prove anything regarding the superiority of one kind of decision above the other. Maybe Singer and Greene reason the other way around: the fact that utilitarian decisions are rational and deontological decisions are emotional shows that reason is superior to emotion. But again, that is question-begging towards old school sentimentalists as much as to philosophers who defend that emotions reveal moral truths (e.g. Roberts 2003; Zagzebski 2003 and my own approach).

This brings us from intuitions to emotions. I already gave an alternative explanation of Greene’s findings from the point of view of ethical intuitionism. Now I wish to add to this by also invoking my specific approach that combines ethical intuitionism with a cognitive theory of emotions (affectual intuitionism). Greene poses a challenge to deontologists: how can our deontological intuitions be true if in fact they are based on emotions? (Greene 2007a, 69). Greene is absolutely right that this is a difficult point for rationalist deontologists. However, my affectual intuitionism does not have any problems with it: the fact that an insight is based on emotions does not preclude the possibility that it is truth-apt. On a cognitive theory of emotions, emotions can be a form of judgement and insight into objective moral truths. In his comment on Greene, Mark Timmons hints in a similar direction, referring to a ‘sentimentalist deontology’ (Timmons 2007). In reply to this point, Greene ends his rejoinder to Timmons with the following words:

Kant was opposed to emotion-based morality because emotions are fickly and contingent in oh-so-many ways … About that, he was right. (Greene 2007b, 117)

It is true that emotions are fallible. But that holds for all our sources of knowledge. Maybe gut reactions are especially notorious, but the moral emotions that I have discussed before involve a high degree of reflection. They also serve as a correction of our immediate responses. But they also correct our overly rational judgements, as those might bias us towards a cold attitude that lets us become cynical and indifferent towards the well-being of others altogether. This becomes clear when we look at another set of moral dilemmas.

In his fMRI studies, Greene also used an example from Peter Unger where in the first case somebody can help a person in front of his eyes who is in need at a slight personal cost, in the second case somebody can donate money to a charity. Most people think we should help the person close by, but we don’t need to make the donation (Greene 2003, 848). However, this means that in this case personal, emotional considerations are superior. The impersonal case is rational but morally questionable. In order to overcome our rational paralysis concerning people who are far away, our emotional—personal judgements should be extended to impersonal cases. Sympathy allows us to care for others who can even be far away.

My approach allows for the following interpretation of Greene’s experiments: emotions such as compassion give access to deontological considerations such as respect for persons. However, in real life people sometimes face genuine dilemmas in which a utilitarian decision is the least unacceptable. Such decisions might involve a
detached calculus that can be provided by rationality. However, the fact that we have to sacrifice innocent people even though we have no choice means that we feel guilt and remorse. But there is nothing irrational about these feelings. They show us and remind us that we had to make a tragic choice, which can help us direct future actions. It would be much more irrational, albeit more pleasant, to not have these feelings.

Now let me make matters more complicated once again. I wonder whether it is really true that utilitarian considerations are purely rational and not emotional. Maybe a utilitarian decision does not need to involve an occurrent emotion, but it could invoke a dispositional emotion (Greene actually makes a similar claim, using different terminology; cf. Greene 2007a, 64). That would explain why on Greene's fMRI scans, emotional brain regions are not highlighted during utilitarian decisions. Nevertheless, this is how dispositional emotions might play a role in utilitarianism: our sympathy for others makes us see the relevance of the utilitarian principle, i.e. that we should improve the well-being of as many people as possible. We do not need to feel sympathy at every instance in which we make a utilitarian judgement: nevertheless, I think that sympathy provides for the underlying justification of the utilitarian principle. However, in concrete decision-making, the very same principle can require us to make a detached judgement, especially if it involves sacrificing the happiness, or even worse, the lives of other people. In such situations, we need to make a fine-grained balance between these kinds of considerations and ‘respect for persons’ considerations that, by being directed at concrete individuals, have a more direct appeal to occurrent emotions.

Look at the following example: I am hungry. I have an urge to eat a greasy snack. However, then I think about my looks and about my health and decide to eat some fresh fruit instead. The typical analysis of a case like this runs as follows: my irrational gut feeling says: go for the unhealthy snack! Then fortunately, reason kicks in and tells me in its authoritative voice to go for the healthy food. A typical example of the famous battle between reason and emotion, System 1 versus System 2 in the terms of DPT. However, I think that this analysis of the example is mistaken. My urge for the greasy food is not an emotion. Urges are affective states, but they are rather primitive compared to full-blown emotions. Furthermore, my decision to go for the healthy fruit is not purely rational and detached. It also involves emotional considerations: I care about my health and my looks. But these emotional considerations are dispositional, they are not felt every time I think about my health and looks. Rather, they predispose me to certain attitudes and ideas. And these dispositional emotions are not only affective states, they also involve cognitive states and a high degree of reflection.

Paul Slovic and his colleagues refer to the ‘dance of affect and reason’ (cf. Finucane, Peters, and Slovic 2003) in judgements about risks, but it might be more subtle and complex: the dance of dispositional and occurrent emotions – both of which also involve (practical) rationality.

6. Conclusion: emotion as a source of moral wisdom about risks

The study of trolley problems can be helpful for understanding ethical considerations about risks. In judging the moral acceptability of risks, we also often face dilemmas: trying to improve the situation for someone can mean introducing risks of harm for someone (another person or the same one). What I have tried to show in this essay is that the fact that we have different intuitions about different kinds of moral dilemmas does not need to mean that there is anything wrong with our intuitions. The cases are
different, and we can point to the factors that distinguish them. Ethical intuitionism provides for a framework that allows that in different situations, different judgements are justified. The fact that our responses to such dilemmas are intuitive or emotional also does not need to make us suspicious. Intuitions and emotions are valuable sources of ethical insight. This also holds for intuitions and emotions about risks (Roeser 2006a, 2007). Purely rational approaches let us overlook important aspects of situations. And every ethical theory needs to invoke intuitions at some point.

These are examples of how moral intuitions and emotions can guide our moral choices when it comes to decisions about risks: compassion with (possible) victims of deeds of others or oneself can let us understand that a risk is unacceptable or has to be diminished. Guilt for victims of our own deeds makes us understand that we did something wrong. This is the case if the wrong consequences were intended, but it can also be the case if they were unintended but foreseen. Contempt and indignation for people who illegitimately impose risks on oneself or others helps us to understand that what they are doing is wrong. These emotions point to the values of autonomy and respect for persons. In all these cases, emotions are not responses to detached, rational moral judgements. As Greene’s studies show, they operate directly. Where Singer, Greene and Haidt think that emotional responses are irrational gut feelings, my approach can help to see that they are cognitive moral emotions that help to tease out illegitimate risk impositions. A different understanding of intuitions and emotions allows for a different interpretation of responses to trolley cases and, more generally, moral judgements about risks. These judgements are not irrational gut reactions, but moral intuitions and emotions that are sources of ethical wisdom that have to be taken seriously.

Acknowledgements
Work on this paper was supported by the Netherlands Organization for Scientific Research (NWO) under grant number 275-20-007. A slightly revised Dutch version of this paper has been published as Roeser, Sabine (2009) ‘Intuities, emoties en onderbuikgevoelens: naar een andere interpretatie van neuroethiek’, Tijdschrift voor toegepaste Arbowetenschap.

Notes
1. In some studies, Greene and colleagues further distinguished between the following kinds of judgements: personal versus impersonal, difficult personal versus easy personal, utilitarian difficult personal versus non-utilitarian difficult personal moral judgements; cf. Greene et al. 2004, 391.
2. On his website, Joshua Greene has an extensive list of trolley problems and related dilemmas in a supplement to Greene et al. (2008): http://mcl.wjh.harvard.edu/materials/Greene-CogLoadSupMats.pdf
3. Haidt (2003) acknowledges this but says that this is the exception rather than the rule – an empirical claim that should be further investigated.
5. A key consideration in the doctrine of double effect is also proportionality (cf. McIntyre 2004), something Hauser does not mention.
6. ‘… that each one is morally bound to regard the good of any other individual as mush as his own …’ (Sidgwick 1874/1901, 380). This principle is roughly similar to the principle of equality that I mentioned in Section 4.
7. With the exception of the monist and utilitarian Sidgwick.
8. A main difference is that in trolley problems, outcomes are certain, whereas in the case of risk, they are uncertain. That does not affect the main point of my argument, but it probably means that moral judgements about risks are more difficult.
References


