

Spring 1-1-2012

On Greene's Neuroscientific Challenge To Deontological Ethics

Ian Martin Tully

University of Colorado at Boulder, ianmtully@gmail.com

Follow this and additional works at: https://scholar.colorado.edu/phil_gradetds



Part of the [Philosophy Commons](#)

Recommended Citation

Tully, Ian Martin, "On Greene's Neuroscientific Challenge To Deontological Ethics" (2012). *Philosophy Graduate Theses & Dissertations*. 22.

https://scholar.colorado.edu/phil_gradetds/22

This Thesis is brought to you for free and open access by Philosophy at CU Scholar. It has been accepted for inclusion in Philosophy Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.

ON GREENE'S NEUROSCIENTIFIC CHALLENGE
TO DEONTOLOGICAL ETHICS

by

IAN MARTIN TULLY

B.A., George Washington University, 2009

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Master of Philosophy
Department of Philosophy

2012

This thesis entitled:
On Greene's Neuroscientific Challenge To Deontological Ethics
written by Ian Martin Tully
has been approved for the Department of Philosophy

Date_____

The final copy of this thesis has been examined by the signatories, and we
Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline.

Tully, Ian Martin (M.A., Philosophy, Department of Philosophy)

On Greene's Neuroscientific Challenge to Deontological Ethics

Thesis directed by Assistant Professor Christopher Heathwood

In this paper, I respond to the case against deontological moral theory that Joshua Greene develops in “The Secret Joke of Kant’s Soul” and elsewhere. Using empirical data he and colleagues collected on peoples’ judgments in various moral dilemmas, Greene attempts to show that deontology rests on unsound foundations. In brief, he contends that the intuitions used to support deontological theory are undermined because they are responses to a morally irrelevant feature he calls “personalness.” I argue that deontologists can respond to Greene’s arguments by drawing a distinction between “practical” and “theoretical” intuitions. I contend that it is only the former sort of intuitions that are undermined by Greene’s evidence, and that deontological theory can be supported purely on the basis of theoretical intuitions.

ACKNOWLEDGEMENTS

The author would like to especially thank Christopher Heathwood, for advising the thesis project, providing very helpful feedback in writing and in person, and for being so supportive throughout the process. Many thanks also go to Eric Chwang, Alastair Norcross, and Brian Talbot for their helpful comments.

CONTENTS

CHAPTER

| | | |
|------|---|----|
| I. | INTRODUCTION | 1 |
| II. | TROLLEY CASES, MORAL DUMBFFOUNDING, AND THE EMPRICAL DATA | 4 |
| | Normative Implications..... | 14 |
| III. | THE NORMATIVE SIGNIFICANCE OF NEUROSCIENCE: RESPONSES TO GREENE AND WHY THEY FAIL..... | 21 |
| | Reasons Good, Emotions Bad..... | 21 |
| | This is Only a Problem for Realist Theories..... | 24 |
| | This is a Caricature of Deontology..... | 25 |
| | Berker's Criticisms..... | 27 |
| IV. | THEORETICAL INTUITIONS AND THE CASE FOR DEONTOLOGY..... | 36 |
| V. | CONCLUSION..... | 53 |
| | REFERENCES..... | 54 |

CHAPTER 1

INTRODUCTION

Philosophers have, in recent years, begun to take a greater interest in empirical work in moral psychology, hoping that this research can help resolve longstanding debates concerning, among other things, the respective contributions of reason and emotion to moral judgment (Nichols 2004; Prinz 2007), the nature of moral motivation and the possibility of a consistent “amoralist” (Nichols 2002), and the justificatory status of moral intuitions (Greene et al. 2001, Greene 2008; Haidt 2001; Singer 2005; Sinnott-Armstrong 2008). While this work is still in its infancy, it has already made important contributions to moral philosophy, injecting new life into stale debates and providing a glimpse of what could become a fruitful new interdisciplinary approach to solving problems in ethics.

Given that this research is in early stages, many of the conclusions drawn from it – both descriptive and normative – remain hotly contested. Some of the most controversial results come from the work of Joshua Greene on the psychology of moral intuitions. In a very influential study conducted in the early 2000’s, Greene and colleagues claimed to discover evidence that moral judgments are produced by two separate systems in the brain. One system is fast, affect-laden, and evolutionarily older: this system, Greene and his colleagues argued, is responsible for the production of our moral intuitions. The second, more recently-evolved

system, is slower, more “cognitive,”¹ and consciously-controlled. It produces more reflective – and, interestingly, more utilitarian – judgments than the “intuitive” system. According to the authors of the study, the two systems compete to produce a moral judgment, with the affective/intuitive system typically winning out over the cognitive when the scenario under consideration possesses certain features which excite our emotions (Greene et al. 2001).

On the basis of the “dual-system” account of moral judgment, Greene, Peter Singer, and others have drawn some striking normative implications. They contend that this research is disastrous for the justificatory status of many of our moral intuitions, particularly those that underpin deontological theories. They urge that we discount the judgments produced by the intuitive/affective system and adopt a more rational utilitarianism.

I argue that this conclusion is ultimately too hasty. I concede that Greene has given good reasons to reject certain deontological intuitions, and I spend a large part of the paper defending him from various criticisms which I think largely misunderstand his argument. Much ink has already been spilled trying to defend deontology from Greene’s attack, and I will argue that most of the current responses are inadequate. Nevertheless, I do not think that Greene has shown that we ought to abandon deontological moral philosophy and become utilitarians. I think that even if certain intuitions – what I will call practical intuitions – are of no epistemic

¹ Greene contrasts “cognitive” and “emotional” in the following way: “‘cognitive’ representations...do not automatically trigger particularly behavioral responses or dispositions, while ‘emotional’ representations do have such automatic effects” (Greene 2008: 40). He adds that “‘cognitive’ processes are especially important for reasoning, planning, manipulating information in working memory, controlling impulses, and ‘higher executive functions’ more generally (40).

value, a committed deontologist could get by with more abstract theoretical intuitions which are very unlikely to be undermined in the way that many of our practical intuitions are.

The paper is organized as follows. First, I outline the key details of Greene's research (and some similar research done by Jonathan Haidt) and then lay out the skeptical conclusions Greene and others draw from this work. I then address various criticisms of Greene's arguments and show that those criticisms are generally inadequate. Finally, I present my preferred response to Greene and argue that it is more robust than other responses to Greene in the literature.

CHAPTER II

TROLLEY CASES, MORAL DUMFOUNDING, AND THE EMPIRICAL DATA

Greene bases his dual-track hypothesis of moral judgment and his revisionary claims about moral philosophy primarily on the results of a study he and his collaborators conducted on moral responses to the “trolley problem.” The trolley problem, as articulated by Foot (1978), and by Thomson (1985), involves a pair of scenarios wherein an out-of-control trolley is speeding toward five innocent people who, for whatever reason, cannot get out of the way in time. In each scenario, you are faced with a choice which, if you make it, will save the five people trapped in the path of the trolley. However, in each case it is inescapable that if you make that choice one other innocent person will have to die. The first scenario I’ll call Switch.

Switch: An out-of-control trolley is speeding down the tracks toward five innocent track workers who cannot get out of the way in time. You are standing near a switch which, if thrown, will divert the trolley on to a side track. However, standing on that side track is a single person who will be killed by the trolley if you throw the switch. Is it permissible to throw the switch?

This scenario is then contrasted with a similar one, called Footbridge.

Footbridge: Again there is an out-of-control trolley speeding toward five innocent workers who cannot escape. Now, however, you are standing on a footbridge over the tracks. Beside you there is a very large man who, if you push him, will fall from the bridge into the path of the trolley. This will stop the trolley, but kill the man. Is it permissible to push the man?

The trolley problem has been of special interest to philosophers because its two scenarios so clearly and reliably provoke contrasting utilitarian and deontological judgments. In the Switch case, most respondents judge that it is permissible – perhaps obligatory – to turn the trolley on to the side track, thereby killing one in order to save five (Hauser, Cushman, Young, Jin, and Mikhail 2007: 4-6).² This is consonant with utilitarianism which enjoins us to choose the action which has the best consequences. One death is bad, but five is five times as bad, so we choose the former. (These are the only two choices we can make.) So far so good. But when presented with the Footbridge case, most people have the opposite reaction: we must not push the large man into the path of the trolley (Hauser et al. 2007: 6).³ This is a characteristically deontological judgment, one which prohibits a certain action despite its good consequences. Though quite similar (five lives for one), the two trolley cases provoke different reactions in the vast majority of people. What gives?

Over the years many philosophers have tried to explain these contrasting judgments. However, the word ‘explain’ is ambiguous in this context. There are two ways we might explain the different trolley judgments. First, we might ask for a psychological explanation. This kind of explanation makes no normative claims on its own, but simply investigates the psychological mechanisms that cause the majority of people to judge Switch one way and Footbridge another. Alternatively,

² The authors conducted a web survey involving “some 5,000 subjects...covering 120 countries, but with a strong bias toward English-speaking nationalities” (5).

³ The specific numbers were as follows: 85% believed it was permissible to turn the trolley; 12% believed it was permissible to push the man off the footbridge (6).

we might ask for an explanation of what justifies peoples' differing judgments. Perhaps there is some deeper principle that provides reasons to flip the switch but not to push. Many times these two kinds of explanation – psychological and justificatory – are run together, but they are conceptually distinct and it is important to keep them apart.

With this distinction in mind, consider one famous example in the philosophical literature which appears to fulfill both explanatory functions. The explanation comes from Judith Jarvis Thomson, who suggests (before later rejecting) that the difference between the two cases hinges on the Kantian distinction between treating people as ends in themselves and treating them purely as means (Thomson 1985). According to Kant, it is never permissible to treat a human being as though he were nothing more than an object, something to be used but not respected (Kant 1785/1998). Thomson uses this principle to explain our differing reactions to the trolley cases, arguing that in the Switch case the innocent person on the side track is not being used to stop the trolley: his death is merely foreseen and he is not a critical part of the plan to stop the trolley. (He is just in the wrong place at the wrong time.) In the Footbridge case, however, the large man is being used to stop the trolley. In pushing him in front of it, we are treating him as though he were an object, not as an autonomous being entitled to respect.⁴

⁴ Rather than talking about means versus ends, we might talk about intending versus foreseeing. The death in the *Switch* case is merely foreseen, but in *Footbridge* it is intended. I don't like this way of carving things up, though. We don't intend the large man's death: his *death* doesn't stop the trolley, his body does, and if he doesn't die but still stops the trolley, we count that as a bonus.

As noted above, there are two ways to read this claim. First, we might read it as a psychological claim: it is our grasp of and adherence to the Kantian prohibition on using people purely as means which causes us to judge the two cases differently. Though this explanation is initially plausible, it runs into trouble rather quickly. For instance, consider a variant of Switch, called Loop (Thomson 1985). In this case, you can again throw a switch and divert the trolley on to a side track, but this time the side track loops back around onto the main track where the five workers are trapped. On the loop track there stands a man who will stop the trolley, but who will die in the process. Notice that in this case, as in the Footbridge case, the man's death is integral to stopping the trolley. However, intuitions here are more mixed. Many people (though fewer than in the Switch case) judge that it is permissible to throw the switch in this case, even though in doing so one is flouting the Kantian prohibition against using people purely as means (Hauser et al. 2007: 6)⁵. So, the story goes, it cannot be the Kantian principle that psychologically explains our intuitions in the various trolley cases. Something else must be going on.

Still, the principle might justify our differing intuitions. Perhaps we ought to treat the two cases differently. However, the Loop case also provides a *prima facie* challenge to the claim that Kant's principle can justify our intuitions, since it seems that many people do not find it to be morally impermissible to use someone purely as a means in all circumstances. If that intuition is robust enough, by which I mean that it is more strongly held, or coheres better with our other moral judgments and

⁵ 56% of respondents, according to the study (6).

principles, then it would seem the principle that we must never use people purely as a means is not true. But of course this is not a knockdown argument, because we might revise our intuition that it is not wrong to flip the switch in the Loop case in light of our stronger adherence to the Kantian principle.

It is precisely this dialectic that has shaped the literature on the trolley problem. Over the years philosophers have applied themselves assiduously to providing a justificatory explanation of our intuitions in trolley cases, creating something of a cottage industry as a result. (See, e.g., (Fischer and Ravizza 1992) for an overview.) More elaborate principles, confounded by ever more elaborate trolley scenarios, have proliferated in a way reminiscent of the literature surrounding the Gettier problems and other areas of philosophy where intuitions about strange cases play a dominant role. Despite all this effort, it is probably safe to say that no definitive solution to the trolley problem has yet been found.

Faced with this state of affairs, some philosophers and psychologists have turned their attentions away from justification and towards psychology. Though psychological explanations and justifications are not the same thing, they are importantly related. Knowing why and how we make moral judgments or form intuitions can shed light on whether those judgments or intuitions are justified. For instance, suppose that I am unwittingly racist, and that my unconscious racial biases cause me to make certain moral judgments. Because my biases are irrational, they cannot provide justification for those judgments. Therefore, knowing that my judgments were the result of racial bias would provide reason to think that those

judgments lacked justification. It would seem then that the psychological origin of my moral judgments plays a crucial role in the truth or justification of those judgments. So, it might turn out that in order to solve the trolley problem, we need to inquire first into the psychology of moral judgment.

Philosophers have long debated the psychological nature of moral judgment, with some, like Kant, emphasizing the role of reason, and others, following Hume, putting the emphasis on the role of emotion. Hume famously claimed that “morality is more properly felt than judged of,” (Hume 1739/2007: 3.1.2), and in the early 2000’s that position gained some empirical support from the work of the moral psychologist Jonathan Haidt. Haidt proposed what he called the “Social Intuitionist Model” (SIM) of moral judgment, which claims that moral judgment is primarily driven by “quick and automatic moral intuitions,” and that “moral reasoning is generally a post hoc construction intended to justify” those automatic responses (Haidt 2001: 13). According to the SIM, a moral intuition “is the sudden appearance in consciousness...of an evaluative feeling” (Haidt and Bjorklund, 2008: 188, my emphasis). This view thus helped to corroborate Hume’s prediction that emotion plays the dominant role in moral judgment, with deliberate, conscious reasoning typically only coming afterward, when we are asked to justify our intuitions to others.

In support of this theory, Haidt cites a phenomenon he dubs “moral dumbfounding.” Haidt and colleagues presented subjects with vignettes concerning “harmless taboo violations,” (e.g., consensual incest, eating a dead pet) and then

asked to make moral judgments about the cases (Haidt, Koller, and Dias 1993). Though many of the subjects quickly judged that the actions in question are morally wrong, they sometimes “[struggled] to find a supporting reason,” and those they did choose were often inadequate to the case (Haidt and Bjorklund: 2008, 197). For instance, they might say that incest is wrong because it will lead to birth defects, but the cases they are presented with render those reasons irrelevant (for instance, they stipulate that the couple use birth control). When this is brought to the subjects’ attention, they largely do not revise their initial judgments. Rather, they simply advert to claiming that “it’s just wrong,” or “people shouldn’t do that,” or else “laughing and confessing they [can] not explain themselves” (Haidt and Bjorklund, 2008: 196-199). Haidt takes this to show that it is not reasons that are driving the subjects’ judgments, but rather emotions. Interestingly, there is a pronounced difference between the judgments of high socioeconomic status (and thus high education) and low socioeconomic status subjects: those in the former group tend not to judge harmless acts to be wrong, while those in the latter group do judge them to be wrong (197). According to Haidt, the lower SES subjects draw from a broader range of moral reasons, which he calls “the ethics of community,” involving “issues of respect, duty, hierarchy, and group obligation,” and “the ethics of divinity,” comprising “purity, sanctity,” and so on (197).

Haidt concludes on the basis of these results and others like them that most moral reasoning is post hoc rationalization for our initial, affect-laden responses. In short, he suspects that we do not hold most of our moral beliefs because we

reasoned our way to them, and we are fairly resistant to changing them even in the face of powerful counter-reasons. While he acknowledges that moral reasoning and moral persuasion do happen – after all, those individuals with more education tend not to judge harmless taboo violations to be morally wrong, possibly (though not conclusively) reflecting the role of moral argument and learning – he expects that this is relatively rare, placing far greater emphasis on culture and conformity to group norms as the primary drivers of our moral lives. We internalize the moral beliefs of our parents and peers without giving those beliefs much in the way of rational scrutiny.

While Haidt's research is fascinating, his conclusions are almost certainly too pessimistic about the role of reason in moral judgment. Building upon and refining Haidt's approach, Joshua Greene and colleagues hypothesized that both reason and emotion play a crucial role in the formation of moral judgments, but that each is, roughly, the product of a separate "system" in the brain. Haidt, then, had only uncovered half the story. While moral intuitions really are the product of a quick, affect-laden system that is resistant to rational correction, there is also a lately-evolved, "cognitive" system that is capable of slow, deliberate reasoning about moral problems (Greene, Sommerville, Nystrom, Darley, and Cohen 2001). These two systems often pull us in different directions, each competing to produce our final moral judgment about a given situation.

To put this hypothesis to the test, Greene and colleagues used fMRI to get a picture of what was going on in the brains of subjects contemplating trolley cases.

First, they asked an independent group to sort a set of moral dilemmas into two groups, “moral-personal,” into which group the Footbridge case went, and “moral-impersonal,” containing, among other cases, the Switch case (Greene et al 2001: 2106). What makes a dilemma personal rather than impersonal? There are three criteria, summarized under the slogan ME-HURT-YOU: “a moral violation is personal if it is (i) likely to cause serious bodily harm, (ii) to a particular person, (iii) in such a way that the harm does not result from the deflection of an existing threat onto a different party” (Greene and Haidt 2002: 519). Moral dilemmas that involve personal violations are personal moral dilemmas. They then presented these cases to subjects and monitored their brain activity using fMRI. What they found was surprising. When subjects were presented with “personal” dilemmas like the Footbridge case, those parts of their brains associated with emotions⁶ were far more active than when they were presented with the Switch case and similar “impersonal” cases (Greene et al 2001). Moreover, “areas associated with working memory”⁷ were “significantly less active” when patients contemplated the personal cases (Greene et al 2001: 2107). Even more striking, for those subjects who decided, contrary to the norm, that it was permissible to push the man in front of the trolley, (or, more generally, for those subjects who had the nonstandard, “utilitarian” response in “personal” cases), it was observed that they took noticeably longer to come to their conclusions (2107). This longer response time was hypothesized to be

⁶ Specifically, “the medial frontal gyrus,” “posterior cingulate gyrus,” and “angular gyrus, bilateral” (Greene et al. 2001: 2107).

⁷ “...the middle frontal gyrus, right” and “parietal lobe, bilateral” (2107).

the result of those subjects “overriding” their emotional responses (Greene 2008: 44). On the basis of this data, Greene and colleagues concluded that “from a psychological point of view, the crucial difference between the [switch] dilemma and the footbridge dilemma lies in the latter’s tendency to engage people’s emotions in a way that the former does not” (2106).

At this point, it is important to note that Greene’s dual-system theory remains only one of many potential psychological explanations of moral judgment, and that his empirical findings have come in for their fair share of criticism.⁸ That being said, for the remainder of this paper I’ll be taking Greene’s psychological story mostly – with one important caveat concerning “personalness” – for granted. If it turns out that the dual-system approach is false, then obviously Greene’s case against deontology goes down with it, but whether it is false or not is not something I am in a position to judge. Nevertheless, it is important to realize that this research is still in early days, so the conclusions of this paper are of necessity somewhat tentative.

To summarize the preceding: from the work of Haidt and others it was concluded that emotions play some significant causal role in the production of moral judgments, but as Greene and colleagues showed, that was only part of the story. It appears that there are two separate systems in the brain, one intuitive and emotional, the other cool and cognitive, which compete to produce moral judgments. Those who judge it impermissible to push the man in front of the trolley do so on

⁸ See, for instance, (Huebner 2011) ; (Klein 2011).

the basis of their quick, emotion-laden intuition; those who do not judge it impermissible must override this intuition, employing their more cognitive faculties. For cases where emotions are less salient – e.g., the Switch case – our higher-level faculties take control and through a process of, roughly, cost-benefit analysis, a more utilitarian judgment results. On the basis of this psychological picture, Greene draws some highly revisionary normative implications, to which I now turn.

1.2 Normative Implications: “The Secret Joke of Kant’s Soul”

At first blush, Greene’s data do not seem so revolutionary. Of course pushing people in front of speeding trolleys provokes our emotional brain into action. How could it not? But Greene insists that this discovery is quite damaging for deontological moral philosophy. “Deontological philosophy,” he claims, “rather than being grounded in moral reasoning, is to a large extent an exercise in moral rationalization” (Greene 2008: 36). In this section, I lay out his arguments for this claim.

First, we should get clearer about the precise nature of “deontological moral philosophy.” Characterizing the various moral theories that fall under the umbrella of “deontology” can be a bit difficult, but there are a few commonalities that stand out. McNaughton and Rawling (2006) highlight three.⁹ First, deontologists emphasize constraints. That is, deontologists claim that there are times when it is impermissible to choose the option which will promote the best consequences. Our

⁹ McNaughton and Rawling also discuss the distinction between agent-relative and agent-neutral reasons, noting that consequentialism only countenances the latter. I leave that distinction aside here.

actions are constrained by various prohibitions – against lying, for example – which we are not to violate whenever doing so will maximize the good. This is not to say, however, that these constraints are always and everywhere inviolable. It might be permissible to lie to the murderer at the door, for instance. Nevertheless, there will be times when it is impermissible to maximize the good. These constraints are often cashed out in terms of rights and duties. Pushing the man in front of the trolley will, *ex hypothesi*, produce the most good, yet a deontologist might claim that he has a right not to be harmed. Constraints are probably the most distinctive feature of deontological moral thinking and they often have great intuitive pull.

Second, McNaughton and Rawling argue that deontology is characterized by a focus on special relationships. According to deontologists, we have certain duties to others that stem from the particular relationships we have with them. Consider, for example, promising. If I make a promise to you, then I acquire a special duty to keep that promise, one that cannot be overridden willy-nilly whenever it will produce more good. Some of our special duties are not voluntary: we have special obligations towards our family and perhaps our friends. Thus, it is right and proper that we favor our children, say, over the destitute elsewhere in the world. (This is not to say that we have no duty of charity, only that placing a greater emphasis on one's kith and kin is not wrong.)

Finally, the deontologist asserts that we have options. Our duty to help others, though stringent, is limited. There is a place for supererogation in deontological thinking. We are not, for instance, called upon to give to charity to the

point of marginal disutility, or perhaps anywhere near that. Doing so is laudable, but not required.

That, at any rate, is how moral philosophers have traditionally understood deontology. According to Greene, however, there is a “hidden essence” to deontological theory that can be uncovered using empirical methods, and that once we reveal the nature of this “essence” we will see that there are good reasons for rejecting deontology (Greene 2008: 37). Deontology, he claims, is the pattern of judgments that results from the exercise of one’s intuitive-affective system; consequentialism the cognitive, rational pattern.¹⁰ The core of Greene’s case against deontology is that our intuitive-affective system – which produces deontological judgments – is compromised in various ways that wind up making deontological intuitions untrustworthy.

The three characteristic principles of deontology enjoy support among a diverse group of people. Nevertheless, Greene thinks they are unjustified because they are built upon or grounded in a pattern of emotional response that is sensitive to what anyone would agree, on rational reflection, is morally irrelevant: what Greene calls “personalness” (Greene 2008: 112). As explained above, “personal” dilemmas – like the Footbridge case -- all exhibit similar features, which Greene summarizes as ME-HURT-YOU. It is moral dilemmas with these “personal” features – and only those -- that provoke our automatic, affect-laden intuitive

¹⁰ As will be obvious, Greene is building his case here on some highly tendentious assumptions in the philosophy of language. As I understand him, his position is that the features of deontology outlined above constitute the *sense* or *primary intension* of the term ‘deontology,’ while the psychological mechanisms he has uncovered constitute the *reference* or *secondary intension*. I find myself pretty doubtful about this proposal, but it is beyond the scope of this paper to challenge it here.

system to respond with a strong negative reaction. As Greene puts it, cases like these set off “alarm bells,” which fairly scream at the agent: DON’T DO IT (Greene 2008: 63). When the case does not conform to the ME-HURT-YOU paradigm, these “alarm bell” emotions do not go off, and we are free to apply rational cost-benefit analysis (as in, e.g., the Switch case).

What’s the trouble with “personalness”? According to Greene, the problem is that whether harm is up-close-and-personal or not is morally irrelevant. It’s the harm that has moral import; the distance between harmer and harmed, say, should make no difference to our moral judgments. Hurting someone does not become more permissible if it is done with a rifle (or with an automated drone) than if it is done with one’s bare hands. Most people, Greene believes, would grant these points on reflection. Yet our psychology has evolved in such a way that it is sensitive to such factors.¹¹ When the harm is done up close, we experience an immediate negative reaction and we recoil: MUST NOT DO THAT, or brains tell us. This powerful negative reaction is what leads moral theorists of a deontological bent to propose constraints on actions – even if the action were to produce the best consequences, we must not do it. Deontologists defend these constraints by claiming that they derive their justification from our natural rights (a person has a right not to be harmed), or from some true moral principle like Kant’s Formula of Humanity as an End. Yet, if Greene is right, these constraints, rather than deriving from any such

¹¹ The story is, roughly, that up close harms are the only kind humans (or proto-humans) were capable of during the era of evolutionary adaptatedness. Thus, we have a sensitivity to those harms that we do not have to other, more distant or “impersonal” harms (flipping a switch, say).

moral principle, are instead rationalizations for a powerful emotional reaction that tracks something – proximity – which is morally irrelevant.

In other words, deontologists grant in many cases the importance of producing the best consequences: how else to explain the widespread intuition that one ought to throw the switch in the Switch case? Yet at times they propose to constrain consequentialism on the basis of rights, or inherent human dignity, or some other principle, arguing that these considerations override the ordinary prescription to produce the most good. What Greene takes his research to show is that, at least in many cases, the deontologist is moved to propose these constraints not on the basis of any rational principles, but rather on the basis of her strong negative reaction to instances of up-close-and-personal harm. But there can't be anything more morally wrong with harm up close as opposed to harm at a distance. So the deontologist's constraints lack justification and ought to be rejected.

A similar story can be told for the deontologist's insistence on options. Consider Peter Singer's famous example in "Famine, Affluence, and Morality" involving a child drowning in a shallow pond (Singer 1972). To save the child, one must wade into the muddy water, ruining one's expensive clothes. Nevertheless, Singer claims – and virtually everyone agrees – we are morally obligated to save the child. But once this point is granted, highly counterintuitive results soon follow. As Singer points out, if we think it is a moral obligation to save the drowning child, then we ought to think it is also incumbent upon us to save children in similarly dire circumstances in foreign countries. It cannot really be denied that we could do

so, through donations to Oxfam and other charities, and yet we choose to spend that money on unnecessary luxuries instead. Moreover, most people seem to feel no compunction about making such a choice. What explains this?

As noted above, deontologists sometimes claim that we have options: that there is a distinction between what duty requires and what is supererogatory. Morality does not demand that we go above and beyond the call of duty, sacrificing our comfortable lives in order to pour money into charitable organizations. If Singer is right, we ought to be giving huge percentages of our income away, far more than almost anyone contributes today. But according to the deontologist, while this is morally laudable, it is not required. Greene, however, sees a different explanation. As he puts it, “It turns out that the same psychological theory that makes sense of the trolley problem can make sense of Singer’s problem” (Greene 2008: 47). That is, what’s driving intuitions in these cases is, once again, emotional salience, which is a function of proximity. We feel obligated to save the child in front of us but not the one overseas because here again our intuitions are tracking the morally irrelevant feature of “personalness.” Because we all agree distance/personalness is irrelevant, these intuitions can be discounted. Thus, not only is the distinctively deontological emphasis on constraints undercut by Greene’s work, but the insistence on options is called into question as well.

Finally, there is the deontological claim that we have duties arising from special relationships. Is this claim undermined by Greene’s work as well? Perhaps not. Deontologists often claim that we have special obligations to our own children,

for instance. One might think that there is a debunking explanation for this intuition along the same lines as the explanations above. But on this point, Greene demurs. He writes:

Taking these arguments seriously, however, threatens to put us on a second slippery slope (in addition to the one leading to altruistic destitution): How far can the empirical debunking of human moral nature go? If science tells me that I love my children more than other children only because they share my genes (Hamilton, 1964), should I feel uneasy about loving them extra? [...] It seems that one who is unwilling to act on human tendencies that have amoral evolutionary causes is ultimately unwilling to be human. Where does one draw the line between correcting the nearsightedness of human moral nature and obliterating it completely? (Greene 2008: 76, my emphasis).

Clearly, then, Greene is hesitant to attack our deeply-held beliefs that our children, and perhaps our other relatives and close friends, deserve special consideration on our part. His reasons are not entirely clear to me, because it is not clear what he means by “be[ing] human” or why he thinks that is important. Moreover, it seems to me that Greene should be very careful here. If he is willing to make an exception for special relationships, then why not make a similar exception for other aspects of deontology? Perhaps Greene simply thinks that our intuitions about the special obligations we owe to our children are so deeply ingrained that we could not be made to act in ways contrary to them, at least not without the application of extreme coercion.

At any rate, on this point it seems deontology may escape unscathed. But that is presumably cold comfort to those, like me, with deontological sympathies. Is there anything to be said in response to Greene’s arguments? I think that there is, though I think what has been said already is mostly missing the mark. I turn now, then, to a discussion of responses to Greene.

CHAPTER III

THE NORMATIVE SIGNIFICANCE OF NEUROSCIENCE: RESPONSES TO GREENE AND WHY THEY FAIL

In this section I will focus primarily on an especially prominent response to Greene by Selim Berker in his paper “The Normative Insignificance of Neuroscience.” Before doing so, however, I want to address some common misunderstandings of Greene’s work and show exactly where they go wrong. Getting these misunderstandings out of the way will help to bring into relief the disagreement between Berker and Greene, and will help to clear the way for the case in favor of deontology that I will mount in section IV.

Misunderstanding #1: Reasons good; emotions bad. On a first pass, one might be tempted to read Greene as making the following claim: deontological intuitions are influenced by emotions in a way that consequentialist intuitions are not, and for this reason they ought to be discounted. This is a common assumption. After all, there are times when our emotions can lead us astray. I might not be able to accurately evaluate the quality of a spouse’s or child’s artistic work, say, because my emotional connection to the person will likely lead me to be far too generous and positive. One might think, then, that Greene is making a similar argument: our emotions are leading us astray from the moral truth. But this is a mistake. First, Greene grants that areas of the brain associated with emotion are active when we

make both deontological and consequentialist judgments; what he argues is that our emotions are much stronger in the former case than in the latter. But it is not the fact that deontological intuitions are more emotional that undermines them; it is, rather, the fact that our stronger (negative) emotional response in dilemmas which generate deontological intuitions is triggered by a feature – “personalness” – that is morally irrelevant. As explained above, Greene claims that certain harms – up close harms – cause stronger emotional reactions, and those reactions lead us to have characteristically deontological intuitions, strong prohibitions despite the consequences. Because distance doesn’t make a moral difference, we should discount these intuitions. In other words, it is not that Greene thinks the fact that it feels really bad (for many people) to contemplate pushing someone in front of a trolley is a good reason to reject the belief that we shouldn’t push people in front of trolleys. It is instead the fact that it feels really bad for no good reason. (Because, as he notes and the data bears out, it doesn’t feel bad to direct trolleys onto innocent track workers by means of throwing a switch.)¹²

Though not a misunderstanding of Greene’s argument, there is a closely-connected argument here that is worth addressing and heading off. One might be tempted to argue that Greene in fact gets things exactly backwards, that his data actually show that “personalness” is highly morally relevant. Though his arguments are unconnected with Greene’s work, the philosopher Michael Slote makes something like this case. Slote defends a version of care ethics which places strong

¹² Again, for *most* people it doesn’t. I myself long felt it *was impermissible* to throw the switch, and there is a relatively stable percentage of people who feel the same. But it is a smallish percentage.

emphasis on a person's motives for doing an action in order to determine whether that action was morally right or not. He believes that "actions are wrong or right depending on whether or not they reflect or exhibit a deficiency of normally or fully empathic caring motivation" (Slote: 2010, 148). Caring is, roughly, an attitude of concern for others, one rooted in empathy (Slote 146-7). Because empathy diminishes as the "salience, conspicuousness, vividness, and immediacy" of a person's plight diminishes (as might happen if that person is overseas and thus far away) Slote contends that we have less obligation towards distant others than we do to those near to us (Slote 149). One might read Greene as merely confirming this viewpoint. We feel far less strongly about people at some distance from us, or people whose plight is less up-close-and-personal, and therefore we have fewer moral obligations towards them.

I do not find this line of argument persuasive. Empathy strikes me as a poor guide to our moral obligations. Suppose, as seems plausible, that ordinary persons feel less empathy for members of other races than they do for members of their own race. This would seem, on Slotes's view, to justify greater moral obligations to members of one's own race (in the same way that, Slote believes, stronger empathy justifies stronger obligations to those nearer to one), something many would find morally repugnant. Similar remarks might apply to biological sex. Since virtually all would agree that differences of race or sex/gender are not morally relevant, even though those differences plausibly track differences in strength of empathy, we should conclude that variation in strength of empathy provides no reason for

thinking that differences in proximity or “personalness” are morally relevant, either.

#2: This is only a problem for “realist” theories. In the course of a discussion of “debunking genealogies” of moral intuitions, like the one provided by Greene, Kelby Mason contends that for the debunking to go through, one must be committed to moral realism and “objectivism” (Mason 2011: 451). She writes, “suppose instead we have a response-dependent view of moral properties, on which for an action to be right...just is for it to be disposed to produce a certain kind of pro- (or con-) attitude in us. [...] Then it looks like it would be hard to debunk moral intuitions” (Mason 451).

Here, one might suppose, is an escape route for the deontologist. She can simply abandon the assumption of moral realism/objectivism, and adopt some form of constructivism, wherein moral facts and properties are “constituted by...the stances, attitudes, conventions or other states or products of human beings” (Levy 2007: 300). Such a view seems highly resistant to the sort of attack Greene mounts. The evolutionary or psychological origins of our moral intuitions can no longer provide reasons for thinking that they lead us astray from the moral truth, since these responses simply create the moral truth.

This, of course, will not fly. First, no constructivist can plausibly claim that we must uncritically accept all our intuitions. After all, many of our intuitions conflict with one another. Second, this response confuses the difference between truth and justification. Whether or not there is any moral truth, certain moral

beliefs can be more or less justified. How? Well, one account of justification that is congenial to a constructivist theory is justification by coherence. On this view, a moral intuition or belief is justified if it “hangs together,” so to speak, with our other beliefs and intuitions. This is congenial to constructivism because there is no reference to any mind-independent moral truth to which our moral beliefs can correspond. But on this view, Greene’s arguments still go through, for what he is claiming is that there is a tension in our web of moral beliefs. We have intuitions in cases that are caused by proximity or “personalness.” We also have the intuition that proximity is not morally relevant.¹³ These intuitions are in conflict, which makes one (or both) of them unjustified on a coherentist view. Thus, Greene’s arguments need not presuppose any robust moral realism. Moreover, and to repeat: it is not the evolutionary origins that undermine deontological intuitions. It’s their sensitivity to an irrelevant feature of cases. This problem affects both realist and constructivist deontologies.

#3: This is a caricature of deontology. Well, yes and no. (This is a point I will return to at some length.) One might complain that Greene ignores examples of reasoning on the basis of deontological principles – consider, for instance, Kant’s (in)famous arguments that one should never lie: a (highly counterintuitive) consequence, or so he thinks, of his abstract moral principles – and insist that his picture of deontological thinking is a strawman which envisions deontologists simply accepting uncritically all of their immediate intuitions about cases.

¹³ Note that there is a prima facie difference between these two kinds of intuition – one practical, the other theoretical. This is a point I will return to at some length in section IV.

Some of this, I think, is on the right track. Greene's portrait of deontology bears little resemblance, at least at first blush, to sophisticated theories like those of Thomas Scanlon or Christine Korsgaard. Now, Greene does not deny that deontologists have developed sophisticated theories to explain and justify their views. What he thinks, though, is that these theories are vitiated by their ground-level input – intuitions about cases. Greene distinguishes two “types” of deontology: “Ground-level deontology,” which “is specifically committed to normative positions that are ‘characteristically deontological,’” and “a metaethical deontological tradition...[whose] aim is to lay out a foundational theory of morality upon which a (nonutilitarian) ground-level theory can be ‘constructed’” (Greene 2008b: 116). What he takes his research to show is that the ground-level variety is in serious justificatory trouble; but, he claims, these problems of justification ‘trickle up,’ so to speak, to infect the more sophisticated theories of thinkers like Rawls and Scanlon. As he puts it, sophisticated deontology faces the “GIGO problem: garbage in, garbage out” (116). Since the inputs – the intuitions about cases – are no good (i.e., “garbage”), they can't be used to build a justified moral theory that retains its deontological character. He sees this as a dilemma: if the intuitions about cases (don't push!) make it through the “rational reflection process” (i.e., reflective equilibrium), then “the so-called ‘moral truth’ now reflects arbitrary features of our evolutionary history” (116). But if these intuitions get weeded out, Greene then claims that we have nothing distinctively deontological left. The upshot is that it looks as though the deontologist is damned if she does, damned if she doesn't.

I will have much more to say about these arguments later on, but for now I'm satisfied to have shown that it is a mistake to think Greene misunderstands the nature of deontological theory, or that he is (simply) attacking a strawman. His arguments are powerful and worth taking seriously by even the most sophisticated deontologist. With this in mind, I turn to a more careful and subtle critique of Greene to see whether it succeeds where the above criticisms failed.

Berker's criticisms. Selim Berker formalizes Greene's argument, which he calls "The Argument From Morally Irrelevant Factors," as follows:

The Argument from Morally Irrelevant Factors:

P1. The emotional processing that gives rise to deontological intuitions responds to factors that make a dilemma personal rather than impersonal.

P2. The factors that make a dilemma personal rather than impersonal are morally irrelevant.

C1. So, the emotional processing that gives rise to deontological intuitions responds to factors that are morally irrelevant.

C2. So, deontological intuitions, unlike consequentialist intuitions, do not have any genuine normative force. (Berker: 2009, 321).

In what follows, I address Berker's important critique of Greene's argument.

Berker makes three key points. First, as his title suggests, he is at pains to argue that "the neuroscientific results are actually doing no work" in Greene's arguments against deontology (Berker: 2009, 294). Once we look beyond the glitzy new technology, we will find that Greene's whole case rests on appeals to certain moral intuitions, the same sorts of intuitions deontologists and utilitarians have been arguing over for decades. Second, Berker argues that it may well not be "personalness" which is driving the different responses to Switch-style and Footbridge-style cases. This claim, if true, would be quite damaging to Greene, because his case against deontology rests on the moral irrelevance of

“personalness.” If that feature is not what causes characteristically deontological intuitions, then there is no reason to think that those intuitions are responding to a morally irrelevant feature, and thus no reason to discount those intuitions. Finally, Berker claims that Greene’s arguments cannot establish the superiority of consequentialist intuitions over deontological intuitions, because he has not ruled out the possibility that “the faculty eliciting consequentialist reactions is also responding to morally irrelevant factors, or failing to respond to morally relevant ones” (Berker 325). I argue that Berker’s first argument fails, that the second, while potentially very serious for Greene, is an as-yet-unresolved empirical matter about which I cannot say very much (other than to provide a sketch of Greene’s recent research on the matter), and that the third misses the point of Greene’s attack, though it does raise important issues to which I will turn in the final section of the paper.¹⁴

Berker’s first criticism is, I think, the least important of the three.¹⁵ It appears to reveal a misunderstanding on Berker’s part of the role of Greene’s empirical data in his normative argument. Berker writes:

So the appeal to neuroscience is a red herring: what’s doing all the work in the argument from morally irrelevant factors is (a) Greene’s identification, from the armchair, of the distinction between dilemmas-eliciting-deontological-reactions and dilemmas-eliciting-consequentialist-reactions with the distinction between personal and impersonal moral dilemmas, and (b) his invocation, from the armchair, of a substantive intuition about what

¹⁴ I should probably note that these arguments are not presented here in the order which Berker presents them in his paper. I’ve changed the order for ease of presentation and stylistic reasons; I don’t think it should have any effect on the arguments.

¹⁵ This is not how Berker himself sees it. In fact, he regards this as his “most pressing worry” (325). In my view, the concerns about “personalness” are much more damning.

sorts of factors out there in the world are and are not morally relevant.
(Berker 326).

Berker is quite right that without “a substantive intuition about what sorts of factors...are and are not morally relevant” no normative conclusion could be drawn from the fMRI data Greene produces. But that, of course, does not show that the empirical data does “no work.” As Greene himself points out, it is not the neuroscience per se that is essential to the argument; rather, the neuroscientific data serves as evidence for Greene’s dual-process theory, and the dual-process theory does do work in the argument (Greene Notes on Berker 2010: 14-15).

Greene’s fMRI data helps to show that deontological and consequentialist judgments are produced by different systems in the brain, and in particular, that the former have a much closer tie to parts of the brain involved in emotions than the latter. Now, with this dual-process picture in hand, which purports to show that there is a crucial difference at the psychological level between deontological and consequentialist judgments, Greene can attack deontological judgments without simultaneously undermining consequentialist judgments as well (Greene Notes: 15). That is a crucial aspect of his argument, and it would not go through without the empirical data. Therefore, Berker is wrong that the neuroscience does no work.

Berker’s second criticism is far more serious. He argues that there is good reason to doubt the truth of P1 – that is, there is good reason to doubt that it is “personalness” that is triggering deontological intuitions in the Footbridge case and cases like it. If that were true, it would be a critical blow to Greene’s argument, for the moral irrelevance of “personalness” does all the normative work in his case

against deontology. If Greene is wrong, then it might turn out that something highly morally relevant is triggering deontological, “constraint-type” intuitions. If that were true, Greene’s argument would fail.

There are several reasons to think that the personal/impersonal distinction might not track the gives-rise-to-deontological-judgment/gives-rise-to-consequentialist-judgment distinction. First, Berker cites Frances Kamm’s “Lazy Susan Case,” which he glosses as follows:

In this case, a runaway trolley is heading toward five innocent people who are seated on a giant lazy Susan. The only way to save the five people is to push the lazy Susan so that it swings the five out of the way; however, doing so will cause the lazy Susan to ram into an innocent bystander, killing him. (Berker 311).

Berker claims that this is an example of a personal dilemma, but he notes that many people – including Kamm herself, have the intuition that it is permissible to push the lazy Susan, thereby killing the man. If that’s right, then it would seem personal dilemmas do not always give rise to deontological judgments, throwing a wrench in Greene’s case. However, I don’t find Berker’s argument here very plausible. For one thing, this is a pretty unusual set of circumstances, to put it mildly. I find it difficult even to picture, and I also find it difficult to see how this case constitutes a personal dilemma: after all, one criterion of such dilemmas is that they typically involve the kinds of harm humans were capable of in the Pleistocene. Surely killing people with giant lazy Susans is not one of those kinds of harms. Moreover, I don’t find that I have a strong intuition one way or another as to the permissibility of pushing the lazy Susan, and I suspect I’m not alone. It could be

that in extremely outré thought experiments, our intuitions go haywire. If so, that would seem to weaken the argument – what Berker needs is a more pedestrian example of a personal dilemma which does not provoke a deontological – constraint-type -- judgment.

Such examples, however, are not difficult to find. Ron Mallon and Shaun Nichols note that “some acts of self-defense, war, and punishment are plausibly personal and emotional, but regarded as permissible nonetheless,” (Nichols and Mallon 2010: 311). They also cite spanking one’s child as a personal, yet permissible, action (311). These appear to be pretty clear counterexamples to Greene’s claim that personal harms trigger deontological, constraint-type judgments of impermissibility.

Greene does have a response to this. In his as yet unpublished “Notes on Berker,” he doubles down on the claim that “personal force” is triggering deontological judgments in trolley cases. In a 2009 study, he and colleagues presented students with four Footbridge scenarios: the original footbridge case and three variants called Footbridge Pole, Footbridge Switch, and Remote Footbridge (Greene Notes: 15-16). Pole differs little from the original case: the only modification is that rather than using one’s hands to push the man, one now uses a pole. In both Footbridge Switch and Remote Footbridge, the man is dropped through a trapdoor in the footbridge onto the tracks below, but in the former case the switch that opens the door is near the man, and in the latter it is far away.

Greene then “asked separate groups of subjects to judge whether the action proposed is morally acceptable,” and got the following results:

standard footbridge: 31% Yes
 footbridge pole: 33% Yes
 footbridge switch: 59% Yes
 remote footbridge: 63% Yes (Greene Notes 15-16).

He notes that the difference in response between Pole and the original case is not significant, and neither is the difference between Switch and Remote, but the difference between the two sets of cases is highly significant (Greene Notes 16).

Clearly, the two sets differ in terms of the degree of personal contact the agent has with the man on the footbridge. Thus, Greene takes this data to show that “personalness” is what’s distinguishing Footbridge and Switch cases.

Is Greene right? Obviously a great deal hangs on this question.

Unfortunately, at this point I think there is little more to be said from the armchair. Whether it turns out that “personalness” or something roughly like it is triggering deontological judgments in trolley cases is an empirical question that clearly needs more attention.¹⁶ Perhaps the cases Mallon and Nichols cite (spanking, self-defense, etc.) have other crucial features that override our native aversion to them, or perhaps Greene’s 2009 results are unusual. At this point there is just not enough evidence to say.

¹⁶ In a 2009 paper, Greene rattles off a list of candidates: “It could be because the harm in that case is more “personal” as defined in Greene et al. (2001), because it’s intentional (Cushman, Young, & Hauser, 2006; Mikhail, 2000; Moore, Clark, & Kane, 2008; Schaich Borg, Hynes, Van Horn, Grafton, & Sinnott-Armstrong, 2006), because it involves an intervention on the victim (Waldmann & Dieterich, 2007), because it is more direct (Moore et al., 2008; Rozman & Baron, 2002), because it involves physical contact (Cushman et al., 2006), because it involves a combination of “personal force” and intention (Greene, Cushman, Stewart, Lowenberg, Nystrom, & Cohen, in press), or for some other reason” (Greene 2009: 6). Obviously, then, more work needs to be done before any firm conclusions can be drawn.

Finally, even if Greene is right about “personalness,” might it not be the case that there is some similar debunking story to tell about consequentialist intuitions? That, in a nutshell, is Berker’s third criticism of Greene. Even if deontological intuitions are undermined by their tight connection with “personalness,” surely that does not show that we all ought to become consequentialists. Berker writes:

Even if P2 does strike us as true, the argument’s conclusion does not follow, for C2 does not follow from C1. Suppose we deem some of the features triggering deontological intuitions to, intuitively, be morally irrelevant, thus granting P2. This is a strike against deontological intuitions. However, we can only conclude that consequentialist intuitions should be privileged over deontological intuitions if a parallel case cannot be made against consequentialist intuitions (Berker 324-5).

Ultimately, and as noted in his first argument against Greene, Berker thinks that Greene’s whole argument boils down to a battle of intuitions about moral relevance. It seems, then, that Greene cannot get the conclusion he wants from his data.

There are two problems with this claim. First, as I’ve argued, Berker is wrong to think that Greene’s empirical data is doing no work. In fact, that data, by revealing an asymmetry in the way that deontological and consequentialist judgments are formed, is providing the main reason to think that deontological, but not consequentialist, judgments are in justificatory trouble. A more important problem, however, is that Berker misunderstands what C2 is claiming. It does not say that we ought to trust our consequentialist intuitions (or judgments – since Greene reserves the term ‘intuition’ solely for deontological responses). All that it says is that we ought not trust our deontological intuitions. These are two different claims. It might be that we ought not trust either consequentialism or deontology.

Perhaps we should be moral nihilists. Perhaps we should be virtue ethicists. So when Berker claims that C2 does not follow from C1, he seems to be reading C2 as the much stronger claim that we ought to trust consequentialist intuitions and not trust deontological intuitions. Of course, Greene himself often appears to endorse Berker's stronger reading of C2, but his more considered view is that he "would rather say [with regard to C2] that the results cast doubt on the reliability of deontological judgments without doing the same for consequentialist judgments" (Greene Notes 13).

All that being said, I do think that Berker is on to something here. It is important to see that Greene has not conclusively shown that consequentialist judgments are not undermined by some heretofore unseen factor, nor has he established the truth of the far more contentious claim that consequentialism is true. In the final section of the paper, I build on these insights, sketching a reply to Greene on behalf of the deontologist.

First, though, it would be useful to summarize the arguments of the preceding section. I have argued that Greene's case against deontology is not so easily refuted as many have supposed. It does not amount to the claim that emotions are bad and reasons good, it is not dodged by "going constructivist," and it is not – at least, not simply – a caricature of deontology which ignores the role of reasoning in deontological moral philosophy. Moreover, though some of Berker's worries are worth taking seriously – particularly his doubts about "personalness" – I do not think he has satisfactorily refuted Greene's claim that we should not trust

our deontological intuitions, and his dismissal of the role of “neuroscience” in Greene’s argument is a serious error. However, Berker is right that Greene has not shown consequentialism to be true, or even shown that we should generally trust our consequentialist judgments. The most Greene has shown, in fact, is that deontological intuitions in certain cases ought not be trusted. This, I argue, is a long way from refuting deontology, as I will now try to show.

CHAPTER IV

THEORETICAL INTUITIONS AND THE CASE FOR DEONTOLOGY

Let us grant that deontological intuitions in the Footbridge case and in cases like it are undermined by their sensitivity to something roughly like “personalness,” even though the empirical evidence is arguably still unsettled on that question. If that is right, then we should not trust those intuitions; they’re unjustified. That’s bad news for deontologists, but all is not lost. The reason is that there are (at least) two kinds of intuition – theoretical and practical – and Greene has only undermined our confidence in practical intuitions.¹⁷ A committed deontologist might be able to make do with theoretical intuitions, a possibility that I will attempt to make more concrete in this last section.

This section is organized as follows. First, I explain in greater detail what precisely I mean by the term ‘intuition,’ and distinguish that term as it is used in philosophy from the slightly different way that it is used in psychology. With that clarified, I distinguish between practical intuitions and theoretical intuitions, and provide reasons for thinking that the latter are more epistemically secure. I then argue that deontologists can make use of theoretical intuitions to support their moral theory, and offer some examples of theoretical, deontological intuitions. Finally, I respond to some criticisms Greene might raise. Importantly, I won’t be

¹⁷ I borrow the terms ‘theoretical intuition’ and ‘practical intuition’ from (Sandberg and Juth 2010).

trying to show that deontology is the true moral theory; rather, I am content to argue that deontologists have resources available to them that are plausibly immune to Greene's arguments. Whether deontology is ultimately to be preferred over its rivals, however, is not something I intend to argue for.

First, what is an intuition? As I'll be using the term, an intuition is an intellectual seeming -- in some ways roughly like an intellectual perception. As Michael Huemer puts it, "intuitions are initial intellectual appearances, that is, states of its seeming to one that something is the case upon intellectual consideration...but prior to reasoning" (Huemer 2008: 370). The "prior to reasoning" bit is crucial. It might seem to me, for instance, that it is wrong to push the man in front of the trolley, but it also might seem to me that five lives are greater in value than one, or that if x is better than y, and y is better than z, then x is better than z, and in none of these cases would that seeming or appearance of truth be based on any prior reasoning. All of these seemings would, on my view, count as intuitions. As should by now be clear, this way of understanding "intuition" is a departure from how Greene understands the term. For Greene, there are two ways of using the term: the philosopher's way, and the psychologist's way. Philosophers use "intuition" in roughly the way that I will use the term. Psychologists, on the other hand, reserve the term "intuition" for "the output of a process to which the individual has no conscious access" (Greene Notes 19). As an example, he cites our ability to sort faces as male or female (19). We don't know how we do it; we just do it. (And effectively, too.) On this account, the snap judgment that one ought to push

the man in front of the trolley is an intuition, because it is fast and opaque to conscious access.¹⁸ Our judgment that it is permissible to throw the switch, on the other hand, is not an intuition. In that case our reasons are clear to us: five lives are worth more than one, so throw the switch.

Is there a way in which we can reconcile the philosopher's account of intuitions with the psychologist's? Yes. In fact, they are not so very different at all. On both accounts, what is most distinctive of an intuition is its non-inferential basis of justification. Thus, what we should say here is that the judgment that one is allowed to throw the switch is not an intuition on either the philosopher's view or the psychologist's, because it is inferential. It is based on reasons, and we are able to consciously introspect and attend to those reasons. We can also conclude that the judgment that one must not push the man is an intuition on either discipline's account. The judgment is non-inferential, as is made clear by the fact that we have difficulty providing any reasons for it and must engage in long philosophical investigation to find some suitable rationale. So, despite the fact that the two disciplines use the term 'intuition' in slightly different ways, philosophers can happily accept Greene's psychologically-motivated claim that the judgment that one should not push the man is an intuition and the judgment that one ought to throw the switch is not.

Now, Greene often seems to believe that consequentialists do not need to rely on moral intuitions. For instance, he claims that "consequentialists consciously

¹⁸ By this Greene means that we judge that it is wrong to push, but we don't know *why* it's wrong. The long search for a rationale that covers various trolley problems attests to this fact.

make their judgments based on the body count and say ‘to hell with everything else.’ If their intuitions tell them otherwise, they ‘bite the bullet’ and say ‘so much the worse for our intuitions’” (Greene Notes 15 – my emphasis). However, I think this is an impoverished account of consequentialist moral thinking. The ‘body count’ all by itself cannot get us to any normative conclusions; in order to make the judgment that one must throw the switch, one needs some normative premises. One might need, for instance, the claim that the death of an innocent person is a bad thing (among other claims). But where do these premises get their justification? Greene doesn’t say, but it seems to clear to me that the reasons upon which the consequentialist judgment that one ought to throw the switch is based – e.g., that pain/suffering is bad, that one ought to minimize harm -- are intuitions. These reasons do not seem inferential; if asked to justify them, one is likely to be met with an incredulous stare. But, and this is important, they are not like the intuition that one must not push the man. They are not intuitions about actions or cases; rather, they are intuitions about principles.

I will be calling these kinds of intuitions “theoretical intuitions.” What I want to investigate now is whether these theoretical intuitions, which in the trolley problem at least appear to support consequentialist thinking, always support consequentialism. I think there are times when our intuitions about abstract principles lend support for deontological theories over utilitarian or consequentialist ones.

Following Huemer and others, I contend that we can have intuitions at different levels of abstraction. Often, we have intuitions about concrete cases. Call these practical intuitions. The intuition that one ought not push the man, for instance, is a practical intuition. There might also be mid-level intuitions, which according to Huemer are “intuitions about principles of an intermediate degree of generality” (Huemer 2008: 383). Finally, there are theoretical intuitions, which are intuitions about abstract principles. For instance, intuitions about the transitivity of “better than” might fall into this category. (For the remainder of the paper I’ll leave mid-level intuitions aside to avoid complicating matters unnecessarily, and will only speak of intuitions about cases and intuitions about principles.)

The preceding distinction is admittedly vague. One would like for there to be specific criteria that an intuition must meet before it can be classified as theoretical or practical. Unfortunately, I confess I don’t have any well-worked out story to tell here. Certain criteria come readily to mind: for example, a theoretical intuition should be an intuition that is not generated by a specific set of circumstances (e.g., a speeding trolley with trapped workers, a drowning child in a shallow pond); there should be nothing particular that one must do (e.g., flip a switch, or push a man); and no individuals should be named or described in any detail. Mostly, though, I simply hope that readers will grant the rough distinction between intuitions about concrete cases, circumstances that are spelled out in some detail, and intuitions about principles, where no specific set of circumstances is referenced, even sketchily. I can, however, offer some intriguing experimental data that helps

provide some evidence for this rough-and-ready distinction. Consider, for example, work done by Joshua Knobe and Shaun Nichols on intuitions concerning moral responsibility. Nichols and Knobe found variation in subjects' intuitive responses regarding moral responsibility which depended on whether the questions they were asked were "abstract" or "concrete" (Nichols and Knobe 2007). In the "concrete" case, subjects were presented with extensively described vignettes involving named individuals and specific moral transgressions and then asked whether the individuals were morally responsible for what they did (having also been told that the individuals lived in a fully deterministic universe); in the "abstract" case, by contrast, subjects were presented with bald assertions of philosophical claims (e.g., "In [a fully deterministic universe] is it possible for a person to be fully morally responsible for their actions?") (Nichols and Knobe 2007: 12). They found that subjects were significantly more willing to attribute responsibility in the concrete cases than in the abstract cases. I think that these results provide some evidence for thinking that there is an important difference between intuitions about cases and intuitions about principles, but I again grant that the distinction is rough and requires additional investigation.

If we do grant this distinction, what follows? Well, perhaps certain of our practical intuitions are untrustworthy. Now, it is not clear that Greene has given good reason to be skeptical about all our practical intuitions, but it might be that at least a large portion of our practical intuitions are not worth very much epistemically-speaking. This could be true if the cases we are asked to evaluate are

often replete with what we might call “confounding” or “undermining” factors, morally irrelevant aspects of the case to which our intuitions are, unbeknownst to us, responding. Thus, we might have good reason to be skeptical about many of our practical intuitions, given that we are rarely certain that we are not responding to what is, on reflection, a morally irrelevant feature. And we are taking it as given that Greene has shown that the intuition not to push in the Footbridge case is undermined and so epistemically worthless. But Greene has given no reason to believe that our more theoretical intuitions are untrustworthy or undermined; in fact, he has committed himself to just the opposite view. That is because Greene’s claim that “personalness is morally irrelevant” is an example, I contend, of a theoretical moral intuition. If those sorts of intuitions were also epistemically suspect, then it is not clear how Greene’s argument could go through. We would be faced with a conflict between two useless intuitions – ‘don’t push,’ and ‘personalness doesn’t matter’ – with no particular reason to prefer one over the other.

Furthermore, as Huemer points out, there are other positive reasons to think that theoretical intuitions are more trustworthy. He notes that they are far less likely to be attended by strong emotions, that they are unlikely to have been “influenced by biological programming” (as, allegedly, in the case of the Footbridge intuition), and that they are less likely to have been influenced by cultural factors (Huemer 2008: 383-4). So it seems we have no particular reason to doubt our abstract intuitions, at least as a class.

Why is this good news for deontologists? Well, according to Greene deontology lives and dies by its (practical) intuitions: a deontologist, says Greene, “cannot dismiss his intuitions as generally invalid because they are his guiding lights” (Greene Notes 19). As noted above, Greene seems to think consequentialists can get by without these intuitions. The reason appears to be that the core idea of consequentialism is so uncontroversial that it needs no support from intuitions about cases: he writes, “everyone agrees that, if all else is equal, it’s better to produce better overall consequences” (Greene Notes 18). What is controversial is the truth of consequentialism, which is the claim that consequences are all that matters. Opposition to this latter claim, Greene seems to believe, comes primarily from intuitions about cases – i.e., practical intuitions -- and if those are (all or mostly) suspect, then consequentialism is presumptively triumphant. But practical intuitions are not the only reasons motivating opposition to consequentialism. There are also theoretical intuitions which support deontological constraints, options, and duties of special relationship. These intuitions are much less likely to be undermined in the same way as the intuition in the Footbridge case is, and therefore Greene and other consequentialists cannot easily discount them.

Some theoretical intuitions only tell against certain forms of consequentialism. For instance, intuitions about fairness often prove troubling for utilitarianism to accommodate. Consider the basic intuition motivating contemporary forms of Luck Egalitarianism, for example:

LE: It is unfair when people are worse off than others through no fault of their own.

This intuition, I submit, has strong initial plausibility. It is not an intuition about any particular case, and I can see no obvious reason why it would be undermined by a morally irrelevant feature like “personalness.” It doesn’t make reference to any particular act one must do, and no individuals are named or described in any detail. It is also an intuition that is in prima facie conflict with certain forms of consequentialism. On a simple utilitarian picture, distribution only matters in a derivative sense. That is, distribution might matter if a more equal distribution, say, would produce greater aggregate welfare, perhaps by assuaging envy. But many people, myself included, have intuitions about fairness that are stronger than this. We think that fairness matters irrespective of consequences. In other words, it might be that a highly unequal society would be best just considering total welfare, yet those of us with strong intuitions about fairness would still claim that there is a moral reason to redistribute, even if this will somewhat reduce total welfare. Of course, fetishizing equality has its limits; we cannot completely ignore the consequences. The point is simply that fairness provides moral reasons independent of consequences for aggregate welfare. That is a controversial view, I admit. But it at least does not seem susceptible to the kind of critique Greene mounts against intuitions in Footbridge cases.

Of course, axiological intuitions about the value of fairness or desert are only problematic for classical utilitarianism; a consequentialism with an expanded

axiology could accommodate them. In fact, some contemporary moral philosophers have attempted to do just that, constructing distribution-sensitive or desert-adjusted consequentialist theories (e.g., Temkin 1993 , Feldman 1997). My point here is not that no form of consequentialism could accommodate desert intuitions and justice/fairness intuitions, but rather that these more sophisticated consequentialist theories provide compelling examples of a moral theory being modified in light of theoretical intuitions.

Greene could grant this much, and yet try to insist that consequentialism remains unscathed, but I think that there are theoretical intuitions that tell against consequentialism as well. Recall that deontological moral theory encompasses three key claims: first, that there are sometimes constraints on action such that it is impermissible to choose the action which produces the most good; second, that people have special obligations to certain others (their families, those they've made promises to, etc.); lastly, that people sometimes have the option to choose actions which produce less than maximal good, or in other words, that some actions are supererogatory. Consequentialism is committed to the denial of each of these claims. As we have seen, according to Greene each of these three legs of deontological theory is supported by our practical intuitions, as when our intuition that we shouldn't push the man in Footbridge lends support to the deontologist's assertion that there are constraints on action. Now, Greene thinks that if you take away the epistemic respectability of those intuitions, you thereby undermine the case for the various aspects of deontology, but I contend that this is only the case if

there are no alternative, more epistemically-secure intuitions which the deontologist can fall back on. Fortunately for deontology, I think that there are some such intuitions.

There are, I submit, a number of theoretical intuitions which provide strong prima facie support for the various aspects of deontology. We might start with the various “prima facie duties” described in W. D. Ross’s *The Right and the Good*. Ross defends the view that there are a number of general obligations – of beneficence, reparation, justice, fidelity, and so on -- which moral agents are compelled to respect, and that these are not reducible to one another or explained by any deeper principle (Ross 1930/2002). These various duties serve as the criteria with which we judge actions to be morally right or wrong: so, for instance, that an action is an instance of beneficence is a mark in its favor – a right-making feature – while the fact that an action is an instance of lying is a wrong-making feature. Importantly, Ross claims that we can know that the prima facie duties are true through intuition alone. The truth of the various claims of prima facie duty is “self-evident”: they are propositions “that cannot be proved, but that just as certainly need no proof” (Ross 1930/2002: 29-30). However, whether any given act is truly right or wrong (and so truly our duty either to do or to abstain from) is far from self-evident. It is a matter of weighing the various considerations (‘is a lie’; ‘is an act of beneficence’) which never results in complete certainty. This last point bears repeating, for it echoes a main contention of this section of the paper. Ross claims that “our judgments about our actual duty in concrete situations have none of the certainty which attaches to

our recognition of the general principles of duty” (Ross 30). He thus claims that certain moral principles are self-evidently true, and that their truth is grasped through intuition. This accords well with my claim that theoretical intuitions are generally epistemically secure. However, he grants that our duty in any particular case is far more difficult to discern, which corroborates my contention that our practical intuitions are often epistemically suspect.

What is important for our purposes here is that several of these duties are in conflict with consequentialism. For example, consider what Ross calls our “duties of fidelity” (Ross 21). These are the duties we have to keep our promises and to not tell lies (which Ross regards as the breaking of “an implicit promise”) (Ross 21). Ross insists that the following principle is self-evident: that the fact that an act is an instance of lying or of promise-breaking is a mark against it, morally speaking. This is not to say that one must never tell a lie or break a promise, for the intrinsic wrongness of such things can be outweighed by other considerations – for instance, by a very strong duty of beneficence – but it is to say that we ought not tell a lie or break a promise whenever doing so we will produce some slightly greater benefit for someone else. As he puts it, “prima facie it is our duty to fulfill a promise irrespective of the precise good of the total consequences” (Ross 1930/2002: 34). This is clearly in conflict with consequentialism, which enjoins us to always choose the action which will produce the most good. Now, I am not suggesting that a consequentialist-minded moral philosopher has no response to this line of argument; all I want to show is that there are certain theoretical intuitions -- in this

case the intuition that an act's being an instance of promise-breaking is a mark against it, morally speaking – which provide prima facie support for deontological moral theory. In this case, the theoretical intuition concerning the intrinsic wrongness of promise-breaking provides support for deontological constraints.

There isn't the space to provide an exhaustive list of theoretical intuitions that are friendly to deontology, but some important examples include intuitions about supererogation and intuitions about justice and innocence. As regards the former, consider the following intuition:

S: It is sometimes permissible to favor oneself or one's family or friends even if in doing so one foregoes doing some other action which would produce more good.

I would argue that S is intuitively quite plausible. It seems reasonable to me to believe that one is not always and everywhere required to produce the maximal amount of good that one can. If this is right, then perhaps the deontologist can defend the existence of options as well. Next consider the following claim from Russ Shafer-Landau:

“[I]t seems to me self-evident that...it is wrong to prosecute and punish those known to be innocent” (Shafer-Landau 2003: 248).

Again, I think this is a principle that would be regarded as highly plausible by a large number of people. If so, and if it is not undermined by some morally irrelevant factor like “personalness,” then I think we have another example of a theoretical

intuition that supports deontology. And there are surely others (for instance, intuitions about consent, or about partiality).

At this point Greene might insist that these are not intuitions about principles at all. If so, I would like to see some principled argument for that claim. These intuitions are not prompted by any thought experiment like the trolley cases, and they can be entertained without reference to any particular set of circumstances. Maybe then Greene would want to say that these intuitions only get their plausibility from their implications about certain cases. This would be a problem for my account, for it would mean that the epistemic status of theoretical intuitions would not be independent of practical intuitions: we would have to check the former against the latter in order to see whether they were plausible or justified. But if that were right, then it would be the intuitiveness of the cases which would be providing support for the principles, and since I grant that Greene has shown that we should not trust our intuitions about many of the cases which support deontology, we would similarly have no reason to trust our intuitions about the principles those cases support. (In other words, we'd be facing the GIGO problem.)

It may be true that many principles are only intuitively compelling given their implications for various cases. However, I suspect it is not always true. That is, I suspect there are many theoretical intuitions which retain their plausibility irrespective of their implications for various cases. Here a test proposed by Huemer may help. In order to tell whether a theoretical intuition has independent

justification – independent of cases, I mean -- we should consider our reactions upon discovering a purported counterexample to the principle which we find intuitively compelling. When we uncover a counterexample to our principle, do we simply revise or abandon the principle, or do we regard the existence of the counterexample as ‘paradoxical’? If the latter – i.e., if we regard it as unthinkable that we might abandon the principle – then we have good evidence that we are dealing with a truly self-evident principle, and so good evidence for trusting the theoretical intuition in question (Huemer 2008: 19). Huemer believes that it is only a subcategory of theoretical intuitions which he calls “formal intuitions” which have this unusual degree of trustworthiness (Huemer 19-20). I agree that formal intuitions (e.g., intuitions concerning the transitivity of ‘better than’) are unusually strong, but I also suspect there are many theoretical deontological intuitions which share something close to this level of trustworthiness. One possible candidate is the intuition above concerning supererogation. That there might be no difference at all between what is morally required and what is “above and beyond” strikes many as counterintuitive in the extreme, and so-called “demandingness” has long been regarded as one of the strongest reasons against (maximizing) consequentialism. I’m not claiming that people typically find, e.g., Singer’s conclusions in “Famine, Affluence, and Morality,” to be “paradoxical” exactly, but they do seem to cut against the intuitive grain to a remarkable extent. That, I suggest, is at least some evidence for the unusual trustworthiness of theoretical intuitions concerning supererogation.

Therefore, I think that one might construct a plausible deontological theory (or, at least, one might attack consequentialism) on the basis of one's more epistemically-secure theoretical intuitions. Here, then, is the kernel of truth in the charge that Greene caricatures deontology. Deontology is not just a bunch of practical intuitions about cases; it is also supported on the basis of intuitions about principles. But what about the "GIGO problem" that Greene raises for more sophisticated versions of deontology? Recall that Greene presented the problem as a kind of dilemma for deontologists. Either they abandon their intuitions about cases, at which point they have abandoned what makes their theory distinctive, or else they incorporate those intuitions and thereby undermine their theory by basing it on unjustified intuitive grounds. We are now in a position to see how to respond to this dilemma: namely, by rejecting the implicit claim that the deontologist's only inputs are her practical intuitions. Since there are theoretical intuitions that support deontological positions, and since those intuitions aren't "garbage," the deontologist can use those intuitions as her starting points for reflective equilibrium. No garbage in, no garbage out, no problem.

Of course, at this point Greene might insist that our deontological intuitions won't survive the test of reflective equilibrium to make it into our final moral theory, but notice how far this takes him from his initial claim that we can use the tools of empirical moral psychology to reveal the poverty of deontological philosophy. For him to make the more circumspect claim that reflective equilibrium will ultimately tell against theoretical deontological intuitions requires getting into

the moral philosophical trenches and engaging in the long twilight struggle between deontology and consequentialism that has been raging for centuries. Maybe he's right, but it certainly isn't obvious, and it certainly isn't the sort of thing one will uncover on the basis of empirical neuroscience. Far more likely, though, is that there are multiple moral systems that, because they are coherent, can be justified through the process of reflective equilibrium.

Finally, one might think that a consequentialist could get by without any intuitions whatsoever, but I can't see how this could possibly go. Recall that consequentialist criticisms of the Footbridge intuition, and the widespread intuition that one ought to save the drowning child but not give huge sums to charity, rest on the theoretical intuition that proximity or "personalness" is irrelevant. If a consequentialist philosopher abandons all his moral intuitions, I don't see how anything other than moral nihilism follows. In short, a refusal to countenance any moral intuitions is nothing more than a refusal to accept any sort of moral truth or justification.

CHAPTER V

CONCLUSION

The recent developments in empirical moral psychology ought to be met by philosophers with enthusiasm. They hold the promise of a fruitful new interdisciplinary approach to longstanding moral philosophical problems. Nevertheless, we ought not adopt any normative conclusions from this research too hastily. While Joshua Greene's dual-process theory is a highly promising account of the psychology of moral judgment, it does not herald the beginning of the end for deontological moral theory. A committed deontologist can defend her moral philosophy on the basis of epistemically-secure theoretical intuitions, eschewing the practical intuitions in cases like the Footbridge dilemma which have been shown by Greene to be epistemically suspect. Of course, whether those theoretical intuitions ultimately withstand sustained philosophical scrutiny is another matter, though for my part I am optimistic.

REFERENCES

- Berker, Selim. (2009). "The Normative Insignificance of Neuroscience." *Philosophy and Public Affairs* 37 (4):293-329.
- Feldman, Fred. (1997) *Utilitarianism, Hedonism, and Desert: Essays In Moral Philosophy*. New York: Cambridge.
- Fischer, J.M and Ravizza, M. (1992) *Ethics: Problems and Principles*. New York: Holt, Rinehart, and Winston.
- Foot, Philippa. (1978) "The Problem of Abortion and the Doctrine of Double Effect." In *Virtues and Vices, and Other Essays in Moral Philosophy*. Berkeley: University of California Press.
- Greene, J.D. (2008). "The Secret Joke of Kant's Soul." In W. Sinnott-Armstrong (ed.), *Moral Psychology*, Vol. 3. MIT Press.
- Greene, J.D. (2008b) "Reply to Mikhail and Timmons." In W. Sinnott-Armstrong (ed.), *Moral Psychology*, Vol. 3. MIT Press.
- Greene, J.D. (2009) "Dual-process Morality and the Personal/Impersonal Distinction: A Reply to McGuire, Langdon, Coltheart, and Mackenzie." *Journal of Experimental Social Psychology*, Vol. 45 (3), 581-584.
- Greene, J.D. (Ms) "Notes on 'The Normative Insignificance of Neuroscience' by Selim Berker." <http://www.wjh.harvard.edu/~jgreene/GreeneWJH/Greene-Notes-on-Berker-Nov10.pdf>
- Greene, J. and Haidt, J. (2002) "How (and Where) Does Moral Judgment Work?" *Trends in Cognitive Sciences*, 6(12), 517-523.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., & Cohen, J.D. (2001). "An fMRI Investigation of Emotional Engagement in Moral Judgment." *Science*, Vol. 293, Sept. 14, 2001, 2105-2108.
- Haidt, Jonathan. (2001) "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review*. 108, 814-834.

- Haidt, J. and Bjorklund, F. (2008) "Social Intuitionists Answer Six Questions About Moral Psychology." In W. Sinnott-Armstrong (ed.), *Moral Psychology*, Vol. 3. MIT Press.
- Haidt, J., Koller, S., and Dias, M. (1993) "Affect, Culture, and Morality, or Is it Wrong to Eat Your Dog?" *Journal of Personality and Social Psychology*, 65, 613-628.
- Hauser, M., Cushman, F., Young, L., Jin, R., and Mikhail, J. (2007) "A Dissociation Between Moral Judgments and Justifications." *Mind & Language*, Vol. 22 No. 1 February 2007, pp. 1–21.
- Huebner, Bryce. (2011) "Critiquing Empirical Moral Psychology." *Philosophy of the Social Sciences* 41 (1):50-83.
- Huemer, Michael. (2008) "Revisionary Intuitionism." *Social Philosophy and Policy* 25 (1):368-392.
- Hume, David. (1739/2007) *A Treatise of Human Nature*. New York: OUP.
- Kant, Immanuel. (1785/1998) *Critique of Pure Reason*. Guyer, Paul and Allen Wood, eds. New York: Cambridge.
- Klein, Colin. (2011). "The Dual Track Theory of Moral Decision-Making: A Critique of the Neuroimaging Evidence." *Neuroethics* 4 (2):143-162.
- Levy, Neil. (2007) *Neuroethics*. New York: Cambridge.
- Mallon, R. and Nichols, S. (2010) "Rules." In John Michael Doris (ed.), *The Moral Psychology Handbook*. New York: OUP.
- Mason, Kelby. (2011). "Debunking Arguments and the Genealogy of Religion and Morality." *Philosophy Compass* 5 (9):770-778.
- McNaughton, D. and Rawling, J. (2006) "Deontology." In D. Copp (Ed.) *The Oxford Handbook of Ethical Theory*. Oxford: OUP. 424-458.
- Nichols, Shaun. (2002) "How Psychopaths Threaten Moral Rationalism, or, Is it Irrational to be Immoral?" *The Monist*, 85: 285-304.
- Nichols, Shaun. (2004) *Sentimental Rules: On the Natural Foundations of Moral Judgment*. New York: OUP.
- Nichols, S. and Knobe, J. (2007) "Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions." *Noûs* 41 (4):663–685.

- Prinz, Jesse. (2007) *The Emotional Construction of Morals*. New York: OUP.
- Ross, W.D. (1930/2002) *The Right and the Good*. Edited by Philip Stratton-Lake. New York: OUP.
- Sandberg, Joakim and Juth, Niklas. (2011). "Ethics and Intuitions: A Reply to Singer." *Journal of Ethics* 15 (3):209-226.
- Shafer-Landau, Russ. (2003) *Moral Realism: A Defense*. New York: OUP.
- Singer, Peter. (1972) "Famine, Affluence, and Morality." In Steven Cahn and Peter Markie (Eds.), *Ethics: History Theory, and Contemporary Issues*. (2009) New York: OUP. Originally published in *Philosophy and Public Affairs* 1.
- Singer, Peter. (2005) "Ethics and Intuitions." *The Journal of Ethics* 9, 331-352.
- Sinnott-Armstrong, Walter. (2008) "Framing Moral Intuitions." In W. Sinnott-Armstrong (ed.), *Moral Psychology*, Vol. 3. MIT Press.
- Slote, Michael. (2009) "Famine, Affluence, and Empathy." In David Boonin and Graham Oddie (Eds.) *What's Wrong: Applied Ethicists and Their Critics*. New York: OUP.
- Temkin, Larry. (1993) *Inequality*. New York: OUP.
- Thomson, Judith Jarvis. (1985) "The Trolley Problem." In Steven Cahn and Peter Markie (Eds.), *Ethics: History Theory, and Contemporary Issues*. (2009) New York: OUP. Originally published in *The Yale Law Journal* Vol. 94, No. 6.