

Does Neuroscience Undermine Deontological Moral Theory?

Abstract: Joshua Greene has argued that several lines of empirical research, including his own fMRI studies of brain activity during moral decisionmaking, comprise strong evidence against the legitimacy of deontology as a moral theory. This is because, Greene maintains, the empirical studies establish that “typically deontological” moral thinking is driven by prepotent emotional reactions which are not a sound basis for morality in the contemporary world, while “typically consequentialist” thinking is a more reliable moral guide because it is characterized by greater cognitive command and control. In this essay, I argue that Greene does not succeed in drawing a strong statistical or causal connection between prepotent emotional reactions and deontological theory, and so does not undermine the legitimacy of deontological moral theories. The results that Greene relies on from neuroscience and social psychology do not establish his conclusion that consequentialism is superior to deontology.

Joshua Greene's work in neuroscience deserves the considerable attention it has received from moral philosophers. The fMRI studies that Greene and his colleagues conducted have opened a rich discussion of the neural activity associated with different types of moral thinking, and this line of inquiry may well have significant implications for moral theory. However, I think there is good reason to be skeptical of Greene's recent attempt, in "The Secret Joke of Kant's Soul," to draw a particular philosophical conclusion from empirical studies [7].

Greene has argued that several lines of empirical research, including his own fMRI studies of brain activity during moral decisionmaking, comprise strong evidence against the legitimacy of deontology as a moral theory. The fMRI studies, Greene maintains, show that deontological thinking arises from areas of the brain more associated with "emotional" reactions, while utilitarian thinking arises from more "cognitive" areas of the brain like the anterior cingulate cortex and dorsolateral prefrontal cortex. Similarly, social psychology studies of "harmless wrongs" show an emotional or non-cognitive basis for judgments that harmless (but disgusting or offensive) actions are wrong, and other studies show that retributive (and therefore deontological) judgments regarding punishment of wrongdoers have an emotional rather than cognitive basis. Greene argues that all of these lines of research suggest a strong connection between emotional responses and deontological moral judgments, and uses this as evidence that deontological moral theory really, "...essentially, is an attempt to produce rational justifications for emotionally driven moral judgments [7: 39]." Deontological theory is a kind of "post hoc rationalization" of emotional reactions, and it is an unreliable moral guide because those emotional reactions evolved in circumstances that are morally different from, and less complicated than, the circumstances that most of us find ourselves in today. Although consequentialism also is a "philosophical manifestation" of an underlying pattern of neural activity, consequentialism is based on more "cognitive" rather than emotional processes in the brain, and so allows "highly flexible behavior" that is responsive to important moral considerations, instead of reflexive "alarm" reactions that may pull an agent away from clear reflection on the morally significant features of a situation [7: 64].

I will argue that the empirical evidence Greene offers does not support the conclusion that deontological theory is primarily a post hoc rationalization of morally unreliable emotional reactions, or that it is therefore inferior to consequentialist moral theory. Although other researchers have pointed out possible

problems with some of Greene's empirical evidence, my main arguments are philosophical – that even if one grants the empirical points, the anti-deontological conclusions do not follow.¹

Greene's Evidence from Neuroimaging

The fMRI studies that Greene and co-researchers have performed on subjects faced with various hypothetical “moral dilemmas” provide the main inspiration for his anti-deontological position. The main point of Greene's studies is to support a “dual process theory” which identifies two types of neural activity involved in moral judgment -- “emotional processes” which Greene associates with deontological judgments, and more “cognitive” processes which he associates with consequentialist moral judgments. Greene rightly acknowledges that “cognition” often refers to “information processing in general,” so both types of process are “cognitive” in the broadest sense, but he is interested in another, “more restrictive” use of “cognition” which is meant to contrast with “emotion” [7: 40]. Although there is no consensus within neuroscience or cognitive science on the exact criteria for distinguishing emotional processes from the more narrowly defined “cognitive” processes, the working definitions typically emphasize differences in automaticity versus conscious deliberation, or motivationally neutral versus behaviorally valenced representations.² Greene maintains that the neural activity involved in consequentialist judgments is “cognitive” in the narrower sense, in that it is less automatic and more behaviorally neutral [7: 40], and that for this reason it is more flexible and provides a sounder basis for moral theory than the emotional processes associated with deontological theory.

In the initial fMRI study in this line of research, Greene and his colleagues identified two different types of brain processes, and associated each with a particular type of moral thinking [3]. This study did not focus directly on a distinction between deontological and consequentialist theories, but it did focus on many cases in which a deontologist and a consequentialist typically might disagree. The researchers presented subjects with twenty one moral dilemmas they classified as “personal” and nineteen moral dilemmas that they classified

¹ I will mention some empirical studies that call into question Greene's picture of a stark dichotomy between emotive/deontological and cognitive/consequentialist processes involved in moral judgment, especially in the section “Other Doubts About Deontology” below [1, 19, 23, 26]. I am not qualified to explore some potential problems, such as the functions of different areas of the brain [21], or that the fMRI studies on which Greene depends offer an inadequate model of moral decisionmaking because they do not include any computational theory regarding moral perception or intuitions [20]. I will also pass over some less neuroscientific and more philosophical issues, such as whether Greene's argument really counts against all deontology or only rationalist deontology [29].

² For a survey of some background sources on the reason/emotion distinction, see [22].

as “impersonal,” along with twenty non-moral dilemmas.³ The study designated a moral dilemma as “personal” if it involved causing direct, serious bodily harm to a particular person or set of people. Otherwise, if a dilemma involved no serious physical harm, or harm only to unspecified victims, or only required diverting some preexisting threat onto different victims rather than initiating the harm oneself, the dilemma was classified as impersonal. A paradigmatic example of an impersonal case, according to Greene and his co-researchers, is a “trolley case,” in which one must choose whether to hit a switch to divert a runaway trolley away from five victims toward a different, single victim. Although the single victim is killed, the harm is caused only by diverting a pre-existing threat, and is not inflicted in a direct, personal way. In contrast, a “footbridge” case, in which the only way to stop a trolley headed toward five victims is to personally shove an innocent passerby into the path of the trolley is a paradigmatic personal dilemma, because the victim is specific, and the harm is inflicted directly and personally. The fMRI scans, conducted on subjects who were asked to classify an action (such as hitting the switch or pushing the passerby) as appropriate or inappropriate, indicated that when considering the personal cases, subjects showed increased activity in “brain areas associated with emotion” [3: 2106] and decreased activity in areas associated with “working memory” or other “higher cognition” [3: 2106, and 7: 43, respectively]. In contrast, when they considered the impersonal moral dilemmas, the subjects’ patterns of brain activity more closely resembled the pattern of activity when considering non-moral dilemmas, with less emotional activity and more activity in “cognitive” areas of the brain. Subjects’ reaction times also were longer in cases requiring personal harm, which Greene took as evidence that in these cases, there was a conflict between an immediate emotional response and a more cognitive calculation of overall effects, and that it took time to exert cognitive control over the emotional response.

Greene has very recently acknowledged that a statistical reanalysis of the results of this first study shows that the study’s results are dubitable, because the lower reaction times for impersonal cases were largely due to several morally obvious “no conflict” cases, in which there was neither a consequentialist nor deontological rationale for an obviously wrong action [10, responding to 19]. But this first study set the basic framework for later studies that Greene still regards as legitimate evidence for a dual process theory of moral judgment (and so as evidence against deontology). In these later studies, Greene and his colleagues turned from a focus on just impersonal versus personal types of situations to the actual, divergent judgments regarding the moral permissibility of inflicting personal harm to promote better overall consequences.

³ The sets of dilemmas used in [3], [6], and [9], and modified for use in [16], are available online at <https://mcl.www.wjh.harvard.edu/materials/Greene-CogLoadSupMats.pdf>.

Greene and his colleagues' second fMRI study, and another subsequent study, ignored non-moral dilemmas and impersonal moral dilemmas, and instead divided the personal dilemmas into difficult and easy cases, depending on how long it took subjects to reach a judgment on the permissibility of directly causing harm [6, 9]. In the second study [6], subjects faced with the difficult personal cases showed the same levels of activity in "emotion-related" areas of the brain as in easy personal cases, but they showed more activity in areas of the brain that had been previously identified with "abstract reasoning, cognitive conflict, and cognitive control," such as the DLPFC, ACC, inferior parietal lobes, and posterior cingulate cortex. In addition, more activity was observed in these "cognitive coordination and command" areas in difficult cases in which subjects decided that personal moral violations were acceptable than in cases in which they deemed the violations unacceptable. The study interpreted these fMRI results as evidence that in difficult decisions about inflicting personal harm for the greater good, emotional reactions conflict with cognitive calculation of costs and benefits, and that subjects who showed more cognitive activity tended to overcome this emotional reaction and ultimately judge the sacrifice to be acceptable because of its greater benefits. A "broader implication" that the researchers draw from the experiment is that "the social-emotional responses that we've inherited from our primate ancestors...undergird the absolute prohibitions that are central to deontology" while "the 'moral calculus' that defines utilitarianism is made possible by more recently evolved structures in the frontal lobes that support abstract thinking and high-level cognitive control" [6: 398]. This foreshadows the objection to deontological theory that Greene develops more fully in "The Secret Joke of Kant's Soul" [7].

More recent studies seem to add weight to the case for associating deontology with emotion and consequentialism with cognitive control. Among these is a study in which Greene and coresearchers asked subjects to reach judgments on difficult personal dilemmas while they were subjected to cognitive load (the subjects were asked to perform an additional cognitive task of identifying numbers scrolling across the screen) [9]. The result, that it took longer for subjects to reach a positive judgment about "utilitarian" actions of inflicting harm for the greater good, but no longer for them to reach "deontological" judgments that inflicting the harm was inappropriate, was taken as further evidence that cognition selectively supports "utilitarian" thinking. A study by Michael Koenigs and coresearchers adopted Greene's emphasis on difficult personal moral dilemmas, and conducted a study showing that patients with damage to their ventromedial prefrontal cortex (VFMPC), an area of the brain associated with social emotion, are more likely than neurologically normal subjects to approve of inflicting harm for the greater good in cases of difficult moral conflict [16]. They take this to support the idea that moral judgment is a dual process system, consisting of "intuitive/affective" and

“conscious/rational” mechanisms, and that absolute prohibitions on inflicting harm depend on the affective or emotional systems [16: 910]. So, several studies seem to link emotion with “deontological” moral judgments or cognition with “utilitarian” judgments about the permissibility of harming others.

But to move from the results of these fMRI studies to the conclusion that deontological theory is inferior to consequentialism requires further steps.⁴ Greene takes it that the fMRI studies establish a strong correlation between emotion and deontology, and between more strictly “cognitive” control and consequentialism. The “prepotent emotional responses that drive people to disapprove of the personally harmful actions” are “characteristic of deontology, but not of consequentialism,” while “the parts of the brain that exhibit increased activity when people make characteristically consequentialist judgments are those that are most closely associated with higher cognitive functions such as executive control” [7: 46]. Greene regards the argument against deontology as “implicit in the empirical material,” but adds that it is “worth spelling it out” [7: 68]. Two claims need to be added to fill out the anti-deontological argument. One is that the emotional responses associated with deontology are an unsound basis for moral theory, and the second is that deontology as a moral theory is really just a rationalization of this set of emotional responses. Greene enthusiastically defends both of these claims.

Greene’s position in [7] regarding the unreliability of deontological judgments is an extension and development of views originally expressed in [5], that “an improved understanding of where our intuitions come from, both in terms of their proximate neural/psychological bases and their evolutionary histories” will show that some of our moral judgments, namely the consequentialist ones, are “more reliable than others” [5: 848]. The problem with emotional moral responses, according to Greene, can be seen by looking at how they evolved. Our emotional aversions toward causing direct, personal harm to others evolved in the conditions of Pleistocene hunter-gatherers, when direct evolutionary advantages would have been gained by peaceful cooperation with a small set of other humans who shared one’s immediate social set and physical environment. But there were no opportunities to “save the lives of anonymous strangers through modest material sacrifices,” the type of impersonal moral situation that Greene regards as central to consequentialist moral theory [7: 47, 7: 59, 7: 70-72, 7: 75]. So humans evolved strong emotional aversions to inflicting personal harm on others, but no similar emotional reactions to “impersonal” moral situations. The strength of our aversion to inflicting personal harm arises from the fact that these emotional aversions “help individuals spread their genes within a social context,” [7: 59] not because such aversions “reflect deep, rationally discoverable moral truths” [7: 70]. So this

⁴ For convenience in quoting Greene, I will follow his practice of using “utilitarian” and “consequentialist” interchangeably.

“contingent, nonmoral feature of our history” often is “morally irrelevant” [7: 70, 75], and our evolved emotional reactions are “unlikely to track the moral truth.” If deontological theory is based on these reactions, then it is based on morally irrelevant factors, and so is unreliable.

Greene’s second step is to argue that deontological theory is “driven” by such emotional judgments, and in fact “...essentially, is an attempt to produce rational justifications for emotionally driven moral judgments” [7: 39]. Taking it as established by the fMRI studies that “characteristically deontological” judgments are generally the result of emotional processes and “characteristically consequentialist” judgments are generally the result of cognitive processes, Greene goes on to point out the likelihood that deontological theory is a post hoc rationalization of these emotion-driven judgments. It is well-established that “humans are, in general, irrepressible explainers and justifiers of their own behavior,” and that “when people don’t know what they’re doing, they just make up a plausible-sounding story” [7: 61]. There is abundant evidence of the human tendency to construct supposedly rational justifications for their intuitive reactions, even specifically in cases of moral judgment [2, 11, 24]. If we “put two and two together,” that is, if we combine the fact that the judgments of deontological theory are “driven largely by intuitive emotional responses” and the fact that we are “prone to rationalization” of our non-rational behavior, we should conclude that deontological theory is “a kind of moral confabulation” [7: 63]. Deontological theory is really an attempt to dress up emotional reactions that are an unsound basis for moral judgment.

Problems with Greene’s “Rationalization” Argument

Greene’s empirical evidence does not show that emotions, or, more specifically, prepotent “alarm” reactions against inflicting personal harm, predict or track deontological theory.⁵ In this section, I will grant for the sake of argument the most central results of the fMRI studies that Greene cites, namely that more “emotional” and more “cognitive” areas of neural activity play roughly the role that Greene’s favored fMRI studies say they do in subjects’ moral judgments in cases of “personal harm.”⁶ Even granting this, it still is not possible to predict the content of deontological theories from the pattern of emotional reactions that Greene identifies. And without this strong correlation, neuroscience provides no compelling reason to think of deontological theories as less legitimate than consequentialist theories. In this section, I will focus strictly on

⁵ His emphasis on the “alarmlike” nature of the emotional responses is from [7], especially p. 63-65.

⁶ See footnote 1. In the next section, I consider some empirical counterevidence to Greene’s claims.

Greene's explicit claim that emotional reactions predict or track deontological theories. In the next section, I will examine other possible interpretations of Greene's main anti-deontological point.

Greene's claim that deontological theories are basically rationalizations of emotional reactions relies heavily on the supposed "...natural mapping between the content of deontological philosophy and the functional properties of alarmlike emotions" and a similarly "...natural mapping between the content of consequentialist philosophy and the functional properties of 'cognitive' processes" [7: 63-64]. Greene himself says that the bulk of "The Secret Joke of Kant's Soul" is devoted to "identifying a factor, namely emotional response, which predicts deontological judgment" [7: 68]. If the emotional reaction against inflicting personal harm really does track the verdicts of deontological theory closely, then the deontologist is left with a huge coincidence to explain, and it is quite plausible to agree with Greene that there is no "naturalistically respectable explanation" for the coincidence that is as likely as Greene's position that deontological theory is a "post hoc rationalization." Greene offers an analogy: If your friend Alice claims that "her romantic judgments are based on a variety of complicated factors," but you notice that over the course of many years she has actually been attracted only to quite tall men, then you are justified in concluding that she is just rationalizing a simple "height fetish" [7: 67].

But in fact, the emotional reaction to inflicting personal harm identified in the fMRI studies does not strongly predict the verdicts of deontology. Although there is no single, quintessential deontological theory, any actual moral theory must cover a much wider range of cases than just the cases that Greene calls personal dilemmas. Since Greene says that the prepotent emotional responses that his fMRI studies identify are only found in cases involving personal harm, these responses can not "predict" what any theory will say about other cases. And any remotely plausible moral theory, including any plausible deontological theory, would provide some guidance in many of Greene's "impersonal" cases. One of Greene's examples of an impersonal moral case asks whether it would be appropriate to hire a black-market surgeon to kidnap a stranger and carve out one of his eyes in order to give you a transplant. I take it that most deontological theories would agree with the 97% of Greene's subjects who said that this would be morally inappropriate.⁷ Since this was an "impersonal" case associated with more activity in cognitive areas of the brain, the verdict is not based mainly on "emotional" reactions, so it serves as a good example of how deontological theories do not show a pattern of tracking emotional reactions. The main point is not merely to nitpick Greene's dilemmas, although I count at least six of Greene's nineteen impersonal dilemmas in which many deontological theories would regard the impersonal moral violation as unacceptable. The main point is that actual deontological theories (as well as actual

⁷ The percentage is given in the online supplementary materials for [9] (see footnote 3).

consequentialist theories) will cover many more cases than the type Greene calls “personal dilemmas.” This narrow range of cases in fact would be likely to form only a small fraction of the cases that would be covered by any actual deontological theory (regardless of whether the deontological theory is morally plausible in its own right). A partial list of some (plausible or implausible, but fairly widely believed) deontological duties that do not involve personally inflicting direct physical harm might include: providing your child a nutritious diet; refraining from sexual relations with your siblings; keeping promises; praying at specified times; promoting others’ welfare when it can be done easily; not shoplifting items you could easily pay for; and paying your income taxes. Since none of these duties involve the “alarmlike” reaction to potentially inflicting personal physical injury on another human, there does not seem to be any way that such alarmlike personal reactions can predict what any or all deontological theories would say about them. The scope of an actual deontological theory will be much wider than a few cases in which inflicting up close and personal harm will promote the greater good.

Perhaps I am being too strict in identifying the types of emotions that Greene thinks should be called upon, in predicting deontological theory. Greene emphasizes that the categorization of moral violations into “personal” and “impersonal,” which depends on the alarmlike aversion to inflicting direct, personal harm, is meant to be a “first cut,” or just a “preliminary step” toward a more adequate distinction between emotion-based and cognitive-based moral judgments [3: 2107]. It is easy enough to grant that an emotional aversion to inflicting direct harm on others may not be the only emotional reaction involved in moral judgments. Disgust is one well-recognized candidate for an emotion that may motivate some moral judgments, and there may be many others. So it is worth considering whether some large set of emotional reactions, taken collectively, can predict or track deontological theory. If so, then Greene’s criticism of deontology as an unreliable guide to moral truth may still be plausible.

But it does not seem likely. It would be a challenging task to find alarmlike emotional reactions, developed early in human evolutionary history, that directly tend to elicit disapproval of activities like tax evasion, plagiarism, or even lying, although many deontological theories would condemn these actions. Guilt or shame presumably will not do the job, because what is needed, for Greene’s purposes, are emotions that precede and motivate moral judgments, not emotions that depend on prior moral judgments. Moreover, besides the general implausibility of positing prepotent emotional reactions which compete with cognition in judging actions like these, Greene’s own studies purport to show that in a great many cases – all the “impersonal” ones – there are only low levels of activity in emotion-related areas of the brain. Since the main focus of the fMRI

studies is on patterns of neural activity, it appears that emotion, as identified by this neural activity, is inert in most moral judgments that are sanctioned by deontological theory.

In addition, another problem arises if one attempts to revise Greene's basic argument by expanding the range of emotions that allegedly underlie deontological moral theory. The problem is that different emotional reactions almost inevitably will conflict. That is, if a set of emotional responses were identified that actually was large and varied enough to collectively track most deontological judgments, then it is likely that such diverse emotions often would conflict with each other as well. So, in a moral dilemma in which one's professional duties (say, to take adequate care in grading students' assignments) conflict with one's family obligations (to spend enough time with one's children), there would most likely be two different prepotent emotional reactions involved (if deontology really does depend on prepotent emotional reactions). If cognitive control is involved in resolving such a deontological dilemma, then it is clear enough that deontological theory is not just a rationalization of prepotent emotions. Instead it is, at worst, a "rationalization" of a set of cognitively influenced moral judgments, which puts it in much the same boat as consequentialism, by Greene's lights. But even if cognitive control is not involved in these (hypothetical) conflicts of emotional reaction, then it is still not accurate to say that the emotions involved predict or track deontological theory. There are conflicting emotions, so they can not all predict what any given deontological theory will require. Even if it turns out to be true that deontological theories somehow crucially leave out a type of cognition that is central to consequentialism, the claim that emotions predict the results of deontological theory is false.

The picture here of deontological theories as encompassing a wide variety of duties is an accurate picture. Even before getting to any specific example of a deontological theory, moral philosophers would generally acknowledge that no deontological theory focuses only on cases of directly inflicting personal physical harm. For example, even the Ten Commandments, a simple but influential version of deontology, include only one commandment against direct personal violence ("Thou Shalt Not Kill"). The Ten Commandments also include religious prohibitions, a duty of sexual fidelity, duties to one's family, a requirement of honesty (not to bear false witness), and an injunction against stealing. More sophisticated deontological theories follow this trend of including a wide variety of duties, including not only prohibitions of direct physical harm to others, but typically also some requirements of honesty, duties of station (obligations related to one's profession or family), and often even a requirement of some sort of beneficence or a duty to help others. If this last sounds more consequentialist than deontological, it is worth noting that deontologists from Kant to WD Ross have endorsed it [14, 15, 25]. Although the duty to promote human well-being is not the

sole priority of deontological theories, it is a duty according to most of them, which underscores the point that focusing only on cases in which promoting the greater good conflicts with deontological requirements is unlikely to capture the real nature of deontology. Greene's model of deontological theory seems to rely on overemphasizing these few cases of conflict.

A closer look at Greene's picture of the construction of deontological theories is instructive. His model is that deontologists start with intuitions about specific cases, especially cases in which deontology imposes "side constraints" on maximizing overall good results, and then construct a theory to accommodate these specific intuitions. If Greene were correct in claiming that there is a conspicuous mapping of these case-specific intuitions onto deontological theory, then this mapping would provide some reason to accept his overall picture of deontological theory-construction, even though it conflicts with what most deontologists might say they were doing. But in the absence of such a close mapping, there is little reason to reject the prima facie evidence that the motivations for developing or defending any moral theory, including a deontological one, appear to be varied. Although some moral philosophers give great weight to particular cases,⁸ others start from more general considerations. Theoretical elegance and consistency with the non-moral aspects of a philosophical system have been powerful forces shaping the moral theories of philosophers from Plato to Kant, and beyond. Even if a philosopher's motives for constructing or defending a particular moral theory are based on some emotions, they may be non-moral emotions, such as religious zeal, or devotion to a teacher or parent. Without the existence of the suspicious tracking relation between emotional reactions and deontological theory, there is little reason to reject the apparent evidence that many forces influence a philosopher's moral theorizing.

The moral philosophy of Immanuel Kant, whom Greene takes to be an arch-deontologist [7], is a good illustration of how unlikely it is that deontologists are really motivated by emotion-driven intuitions about a few specific cases. Kant's own stated strategy, the relationship between his moral theory and the rest of his critical philosophy, and the general timeline of the development of different aspects of his moral theory, do not lend themselves easily to Greene's "rationalization" account of deontology. In *Groundwork of the Metaphysics of Morals*, Kant specifically disavows, and condemns, the strategy of starting from intuitions about particular cases, an approach that he calls "popular philosophy" [15: 210-211]. Instead, he starts from more general intuitions about the nature or concept of morality (such as that moral requirements are commands, and must apply to everyone) and tries to derive the content of the categorical imperative from these widely accepted conceptual claims. The nature of the Categorical Imperative, or supreme moral principle, is fairly analogous to

⁸ Frances Kamm is a striking example.

the role of the Categories in Kant's theoretical philosophy. Basic moral principles are internal or supplied by each rational being herself, but we can see that they are inescapable because they are necessary preconditions of an unavoidable activity of human reason (the activity of deliberating about what to do). This mirrors the status of the Categories, which Kant argues are rules or organizing concepts that we supply as necessary preconditions for coherent perception and theoretical thinking. The analogy between moral principles and theoretical organizing principles is no coincidence, since Kant says in the *Critique of Practical Reason* that the ultimate aim of his metaphysical and epistemological system is to make room for rational belief in God, freedom, and morality [13: 28-29]. His theoretical philosophy may be, in part, a rationalization designed to support a moral system, but it does not appear that either his theoretical or moral philosophy is a rationalization of intuitions about specific moral dilemmas. It is not until literally years into the development of his overall philosophical system that he begins deriving many specific moral duties. His statement of his strategy in moral philosophy came in 1785 [15], and (although four examples of particular duties appear in that work) his more extensive examination of specific duties comes twelve years later [14]. It is hard to make all this fit a model of deontological theorizing as starting from intuitions about particular cases.

The only reason to reject the prima facie evidence that a wide variety of philosophical and emotional motives drive deontological theory, and to regard deontology instead as a rationalization of prepotent emotional responses to specific cases, would be if there really were a suspicious pattern of matching between the case-specific emotional reactions and the verdicts of a deontological theory. But Greene has not established this kind of suspicious correlation.

Other Doubts About Deontology

It may seem that I must have missed Greene's point. While he sometimes says that his main criticism of deontological theories is that they are just elaborate rationalizations of primitive emotions, and that this criticism depends on the emotions tracking or predicting deontological theories, he also sometimes puts his point differently. He sometimes says that conflicting judgments in cases of inflicting personal harm for the greater good are either "characteristically consequentialist" or "characteristically deontological" [7: 39, 7: 65]. Perhaps he means that these cases of conflict somehow reveal the true nature of consequentialism and deontology, even if they fail to predict all of what a deontological or consequentialist moral theory may require. Or, at minimum, perhaps in such cases of conflict, we should regard "characteristically deontological" moral

judgments as “garbage” [8: 116] and should disregard them, so that when consequentialism and deontology conflict, we should always give greater weight to consequentialism’s verdicts—maybe emotion is a distorting influence in a certain type of case.

But these alternative versions of Greene’s anti-deontological argument are highly problematic, both for philosophical reasons and because they run into contrary empirical evidence from neuroscience and social psychology.

It may seem reasonable, at first glance, to think that difficult cases in which consequentialism and deontology give conflicting verdicts may reveal something about the deep nature of each type of theory. Of course, consequentialism and most deontological theories may coincide in their moral judgments about most cases (killing innocent strangers is wrong, failing to make a minimal effort to save a drowning child is wrong, etc.), but still, the cases in which they conflict may reveal what really “drives” each theory [7:59, 7: 63] . A good analogy, supporting this line of thinking, would be that two political parties may agree on most issues, because the best course of action is obvious, but that more controversial cases of conflict reveal what each party is really about. This picture fits with Greene’s overall defense of a “dual process” account of moral judgment, in which cognitively based calculations of effects and emotionally based deontological restrictions may often coincide, but nevertheless are the products of fundamentally different neural systems which sometimes compete.

But this line of thought, though initially plausible, does not stand up to closer examination. Most glaringly, emotional reactions to inflicting personal harm can not “drive” all deontological judgments, for the reasons described in the previous section. Namely, these emotional reactions are only found, according to Greene’s own studies, in cases of inflicting personal harm, not in the wide variety of other cases encompassed by typical deontological theories. And, again according to Greene’s own studies, it is not a viable strategy to expand the range of relevant prepotent emotional alarm reactions beyond just reactions to inflicting personal harm, because Greene’s studies show low levels of neural activity in emotion-related areas of the brain in many cases (cases of impersonal moral dilemmas) about which typical deontological theories would give clear verdicts. There is no apparent way that emotional reactions can “drive” deontological theory, when many or even most judgments delivered by deontological theory are about cases in which there is no particular emotional alarm reaction involved.

This problem is consistent with some cautions derived from empirical research, about generalizing from specific types of cases to conclusions about human thought processes in general. The specific context and content of scenarios can affect subjects’ approach to them, and limit the legitimacy of drawing general

conclusions [28, 1: 383, 1:394]. When applied to Greene's conclusions, this implies that the conflict between cognitive and emotional processes in cases of difficult personal dilemmas may not be indicative of two competing systems involved in moral judgment more generally. In fact, the emotional reactions in cases of difficult personal dilemmas may be the result of feeling torn or conflicted, rather than being reactions to personal harm violations [18, 28, 22: 102-103]. In a 2006 study, Jana Schaich Borg and her co-researchers specifically suggest, based on their fMRI results from more specifically sub-divided cases of moral dilemmas, that the emotional activity that Greene et. al. identified in cases of personal dilemmas may be a reaction to a conflict of values (promoting greater good versus avoiding direct harmful action) rather than merely a reaction to the possibility of taking direct harmful action [26]. So there is good reason to be skeptical of any claim that Greene's fMRI results reveal what really "drives" deontological judgments in general.

In addition, if Greene is correct that consequentialist approaches to moral judgment are more cognitive and deontological approaches are based more on emotional processes, then we should expect to find that people who in general think more cognitively or have more cognitive capacities will tend to make more consequentialist judgments, while more intuitive people will tend to make deontological judgments. But there is empirical evidence that this is not the case. The most direct evidence is provided by a study that examined the relationship between working memory capacity and decisions about various hypothetical moral scenarios [23]. The researchers identified working memory capacity as an indicator of individuals' "cognitive" abilities, in the form of both "controlling emotion and engaging deliberative processing" [23: 550]. One of their findings was that subjects with higher WMC (more "cognitive" subjects) did not approve of "consequentialist" choices (to inflict direct harm in order to promote a greater good) more often than other subjects.⁹ This counts against a strong connection between cognitive thinking and consequentialist thinking.¹⁰

If we regard longer reaction times (RT) as indicative of more cognitive consideration, and shorter RT as indicative of more automatic or intuitive decision making, as Greene sometimes does,¹¹ then there are further problems for the view that more intuitive or emotional people tend to make more deontological decisions. Greene and his co-researchers [9] classified subjects responding to difficult personal moral dilemmas into high-utilitarian respondents and low-utilitarian respondents, based on how often they deemed it appropriate to inflict

⁹ The only scenarios in which there was a correlation between WMC and choosing to sacrifice some for the sake of others was in cases in which the deaths of the few were inevitable anyway.

¹⁰ Also, a study by David Bartels adds some support for Greene but also finds that more deontological people (defined as people holding personal values) made more consequentialist choices than people without personal values in some circumstances, suggesting that cognitive ability can not be the only predictor of "deontological" versus "consequentialist" choices [1].

¹¹ [6: 397] also see [26: 813]

personal harm for the greater good, and they compared the average RT for each group to give utilitarian or non-utilitarian responses under cognitive load or under normal conditions (with no extra cognitive load). The fastest average RT of all was for high-utilitarian subjects approving of utilitarian choices in the absence of cognitive load. Thus, if RT is an indication of cognitive involvement, then utilitarians making utilitarian judgments were less cognitive than low-utilitarian subjects making either utilitarian or deontological judgments. Of course, there is a possible explanation for this. If it is the case that high-utilitarian subjects have more cognitive ability in general than low-utilitarian subjects, maybe they can employ cognitive control more easily and thus more quickly overcome prepotent emotional responses. In that case, the high-utilitarian subjects' fast RT for approving of consequentialist choices would be compatible with saying their choices involve more cognitive command and control than low-utilitarian subjects' choices. But the study by Moore et. al. finds that this explanation is not viable [23]. Subjects with higher WMC (which the study regarded as an indicator of cognitive thought processes) actually took longer to give "consequentialist" answers than subjects with lower WMC did [23: 556]. In other words, more "cognitively" oriented people took longer, not shorter times to approve of personal moral violations (and they did not end up approving of these "consequentialist" harms at a higher rate than less cognitively oriented subjects). If this is accurate, then Greene's high-utilitarian subjects must not have been relying on cognitive processes to reach their fast utilitarian judgments in difficult personal dilemmas. It is worth noting that Moore et. al. believe that their study shows that the cognitive activity involved in difficult personal dilemmas does not serve just to compete with and override "prepotent emotional responses," but rather that there is a "selectively engaged, voluntary reasoning system" that engages in "deliberative reasoning" to take account of relevant emotional and cognitive factors and reach a moral judgment [23: 556]. In other words, it may well be that people with strong cognitive tendencies do not tend overall to override emotional responses more easily and make more consequentialist choices, but just tend to think more carefully about moral judgments.

One final possible interpretation of Greene's anti-deontological argument deserves attention. Above, I have argued that there is good reason to doubt whether "characteristically deontological" or "characteristically consequentialist" judgments, meaning judgments in which deontology and consequentialism conflict, really reveal what underlies all of deontological or consequentialist moral theory. But if we examine these cases of conflict in their own right, rather than as signs of what "drives" each type of moral theory, then perhaps Greene at least provides reason to doubt the reliability of deontological theories in these particular kinds of situation. That is, maybe emotion selectively distorts deontological judgments only in these situations, so when

consequentialism and deontology conflict, maybe we have reason to trust consequentialist judgments more than deontological ones. Sometimes Greene does seem to be making this argument. He admits that “sometimes” deontological arguments can be “cognitive,” but says that in general they “tend to be reached on the basis of emotional responses” [7: 65]. Greene also maintains, in responding to a criticism by Mark Timmons, that the real problem with characteristically deontological judgments’ influence on deontological theory is the “GIGO,” or “garbage in, garbage out” problem [8: 116]. Greene explains that if our characteristically deontological judgments in cases of conflict are “based on emotional intuitions,” then they taint any process of “rational reflection” about moral theory. So, even if these deontological judgments do not drive all of deontological theory, perhaps they are unreliable in their own right (“garbage”) and should be discounted wherever they are found.

But even this more limited criticism of characteristically deontological judgments runs into serious problems. The main problem is that it is not at all clear that these negative judgments about inflicting personal harm for the sake of promoting the greater good are especially “emotional” or non-cognitive. It depends on which other judgments one compares them to, and they inevitably turn out to be as “cognitive” as many judgments that Greene thinks are reliable consequentialist judgments. This is so whether one uses fMRI results as the standard for how “emotional” or “cognitive” judgments are, or whether one uses RT as the indicator.

Suppose that neural activity in different areas of the brain, as revealed by fMRI testing, is the standard for whether judgments are based on prepotent emotional reaction or on more cognitive assessments. Then are “typically deontological” judgments especially emotional, compared to “typically consequentialist” judgments about difficult cases? They are not. Greene’s own finding is that activity in areas of the brain related to social emotions are similar in all cases of “personal” moral dilemmas, whether the cases are easy or difficult, and whether the subjects in difficult cases approve or disapprove of inflicting harm for the greater good [6: 392, 7: 45]. Greene’s position is not that deontological judgments in difficult cases are the result of higher levels of emotion-related brain activity, but rather that consequentialist judgments in these cases involve more neural activity in “cognitive control and command” related areas of the brain, and that this cognitive control overrides prepotent emotional responses [6, 9]. So, is the problem with the deontological judgments in difficult cases that these judgments do not involve enough cognitive activity? Although Greene’s studies did show less cognitive activity for deontological judgments in tough cases than for consequentialist judgments in tough cases, deontological judgments in these difficult cases nevertheless showed more cognitive activity than in (uncontroversial) moral judgments about easy personal cases, judgments which Greene does not question [6:

396]. So it is not that Greene suggests that deontological judgments in difficult personal cases simply fail to meet some minimum level of cognitive control. Then there is no obvious reason to regard these deontological judgments as especially defective – they involve more cognitive activity than many moral judgments that Greene does not question, and do not involve more activity in emotion-related areas of the brain than consequentialist judgments in difficult cases. This conclusion is only strengthened by one of Schaich Borg’s fMRI results, which shows that deliberation about cases in which equivalent consequences result from either directly inflicting harm or merely allowing harm “invoked increased activity in areas dedicated to high-level cognitive processing and suppressed activity in areas associated with socioemotional processing” [26: 813]. If deontological judgments against inflicting personal harm are based on prepotent emotional responses, then subjects should have shown more activity in areas of the brain related to emotion in cases involving a prospect of directly harming people. All in all, fMRI results do not support the claim that “typically deontological” judgments that it is wrong to inflict personal harm are “garbage.”

Reaction times are another possible indicator of whether a judgment is “cognitive,” as opposed to intuitive or emotional. In fact, automaticity of responses versus consciously accessible deliberation is a common standard for distinguishing emotional versus cognitive responses in psychology [22: 100-101], and this distinction is closer to non-technical “common sense” definitions of thoughtful versus hasty judgments than any definition that relies on neural activity in different brain regions. But the data on RT does not, overall, support the idea that “typically deontological” judgments in tough cases are less cognitive than “typically consequentialist” judgments. One study by Greene and his colleagues does find that cognitive load causes subjects to take longer to reach consequentialist judgments in difficult personal cases but that cognitive load does not affect RT for reaching deontological judgments in these cases [9]. Greene takes this as “evidence for the influence of controlled cognitive processes in moral judgment, and utilitarian moral judgment more specifically” [6: 1144]. But the same study shows that under no cognitive load -- that is, under more common circumstances for moral deliberation -- RT for reaching consequentialist and deontological conclusions in difficult cases was virtually identical, and the result is confirmed by both an earlier study by Greene and a study by other researchers [6, 9, 23: 555]. So if RT is taken as an indicator of how “cognitive,” or consciously accessible, moral judgments are, then it is not at all clear that typically deontological judgments are less cognitive than typically consequentialist judgments. Also supporting the idea that reactions to inflicting personal harm are not especially quick and automatic is a finding by Schaich Borg et al, that there is no difference in RT between cases of subjects considering harming people and harming objects [26: 808]. If harming people elicits

an especially fast and automatic, non-cognitive response, as Greene proposes, then RT for harming people ought to be lower than for damaging objects. Overall, Greene's finding that cognitive load selectively interferes with consequentialist judgments must be weighed against significant RT-based evidence against the claim that deontological judgments are especially automatic and emotional, and therefore unreliable.

One point deserves special emphasis, about the importance of fMRI data versus RT data in identifying moral judgment as more cognitive or more emotional. The point is that RT is much closer to ordinary, non-technical ideas about the reliability of moral judgments. If a moral judgment were "automatic" in the sense of being made very quickly (say in half a second, for example), and not being subject to rational revision through deliberation, then most people would agree that the judgment would be unreliable. Emotional, automatic reactions of moral disgust to practices like ethnically mixed marriages, homosexuality, or stem cell research are paradigm examples. But there is no obvious or uncontroversial connection between very quick, "automatic" moral judgments and judgments that involve some given ratio of neural activity in different areas of the brain. Greene himself says, regarding high-conflict personal dilemmas, that "the RT data raise doubts about moral judgment as unreflective, as our participants routinely exhibited RTs over 10s, and in some cases over 20s..." [6: 397], and his own study shows that RT for utilitarian and non-utilitarian judgments in these cases is virtually identical [9]. So there seems to be little reason to regard "typically deontological" judgments in these cases as especially automatic, unreflective, or unreliable. The possibility that moral judgments involving emotion (especially neural activity in emotion-related areas of the brain) can be "cognitive" in the sense of being deliberative or thoughtful undercuts Greene's sharp dichotomy of emotion-related and cognitive-control-related processes in moral judgment, and it is more consistent with the finding of Moore et. al., that cognitive activity can indicate "deliberative reasoning" about morality instead of "simply" restraining emotional reactions [23: 556].¹²

Greene's Other Evidence

Besides relying on fMRI studies, Greene cites other evidence for the claim that deontology is based on emotion and is therefore less sound than consequentialism. The main additional arguments are based on social

¹² It also is consistent with a conclusion of [26] that "In contrast to the speculations of Greene, Nystrom, et al.(2004) and Greene, Sommerville, et al. (2001), our data suggest that some deontological responses (such as the DDA-implied intuition to refrain from action) can be mediated by reason (as suggested by Kant and, more recently, by Hauser), whereas other deontological responses (such as the DDE-implied intuition to refrain from intentional harm) can be mediated by emotion (as suggested by Hume and, more recently, by Haidt)" [26: 815].

psychology studies of two topics: reactions to cases of “harmless wrongs,” and subjects’ intuitions on retributive versus consequentialist approaches to punishment of criminals [7: 50-58].

The study of harmless wrongs on which Greene primarily relies is a study by Haidt, Koller, and Dias [12], which presented a number of cases of “harmless wrongs” to subjects in Brazil (in an affluent city, Porto Alegre and a poor city, Recife) and in Philadelphia, and asked them to answer questions about the scenarios, including “Is it very wrong, a little wrong, or is it perfectly OK for [specific act description]?” The scenarios included actions that were offensive or disgusting, but did not harm anyone, such as a man masturbating into a chicken carcass before cooking and eating it, a son breaking his promise to his dying mother to visit her grave, a woman cleaning her toilet with the national flag, and a family eating its dog. The study found that the responses varied depending on location (fewer respondents in Philadelphia than in Recife thought the actions were wrong), age (adults were less likely than children to regard the actions as wrong) and socioeconomic status (high SES subjects were less likely to regard the actions as wrong). In order to use these results as support for his position, Greene adds two additional claims. First, he suggests that each of the variables (age, location, and SES) is related to how “cognitive” the respondents are – the older, the more “Westernized,” and the higher SES a subject is, the more likely she is to approach moral issues in a cognitive manner [7: 56-7]. Second, Greene connects judging a harmless but offensive action wrong to deontology, and judging that such an action is acceptable to consequentialism [7: 57]. Putting these ingredients together, Greene concludes that the study shows that more cognitively oriented subjects tended to make consequentist judgments, supporting his overall position that consequentialist judgments are more cognitive, and thus more reliable, than deontological judgments.

But the study Greene cites does not support either of the two premises Greene adds in order to reach his anti-deontological conclusion. Instead, it is more or less at odds with both of Greene’s additions.

Far from endorsing an idea that degree of westernization or SES are any kind of absolute, reliable indicators of more cognitive (as opposed to emotional) approaches to morality, the Haidt study instead questions the universality of the then-standard “Cognitive-Developmental View” proposed by Piaget and Kohlberg, which “limited the domain of morality to actions that affect the material or psychological well-being of others” [12: 614]. In opposition to this view that harm to others provides a more cognitively developed, transcultural standard for the realm of moral concern., Haidt et al conclude in their “discussion” section that “the relationships among moral judgment, harm, and affective reactions may be culturally variable” (12: 625). The basic point of the study is to show that previous studies, which were performed only on subjects in westernized

countries, overemphasize one dimension of moral concern, namely harm to others, and underemphasize two other dimensions that are common in other cultures, namely “the ethics of community” (which has to do with a person’s “social role”) and the “ethics of divinity” (which has to do with a spiritual attempt to “avoid pollution and attain purity and sanctity”) [12: 614]. They do not conclude that an exclusive concern with harm or consequences is a sign of a more cognitive approach to moral judgment, but rather that cultural influences shape one’s conception of the realm of moral concern. They certainly do not endorse any idea that a more westernized approach is more cognitive or reliable. Neither does the study suggest that higher-SES subjects have a more cognitive or reliable approach to moral judgment. Of course, there may well be some senses in which high SES contributes to cognitive development. Good nutrition is a necessary condition for maximal brain development, and high-SES children are likely to have access to more formal education, so they may well possess some skills and some types of knowledge to a greater degree than people of low SES. But that is not to say that they are likely to exhibit fundamentally different, less cognitive, neural processes than high-SES subjects. If Greene means to maintain that income and social status play a large role in determining the neural processes involved in moral judgment, then this is at the very least a highly controversial thesis, touching on longstanding debates about nature versus nurture in a particularly volatile way. It is at least as plausible to accept the spirit of the actual study [12], that in making moral judgments, differences in the amount of emphasis on harm versus on social roles or feelings of disgust are largely a product of cultural influences, instead of a measure of how “cognitive” subjects’ moral judgments are according to some absolute standard.¹³ All in all, Greene’s attempt to correlate a more cognitive approach to moral judgment with westernization, or with high socioeconomic standing, is at odds with the study on which he mainly relies.

It is similarly problematic for Greene to append to the study any claim that judging harmless actions to be morally acceptable relies on a fundamentally “consequentialist” approach, or that judging them wrong relies on a deontological approach. Greene says, “In this study, the connection between a reluctance to condemn and consequentialism is fairly straightforward” [7: 57]. But it is not. Haidt et. al. do sometimes describe the traditional Kohlbergian approach to defining morality as depending on “personal harmful consequences” or on “acts that have ‘intrinsically harmful’ consequences to others” [12: 614]. But they clarify, on the same page, that

¹³ Haidt et al’s suggestion that it is misleading to use college students as subjects, because they tend to share a relatively high SES and an emphasis on harm-based standards of morality, is equally amenable to being interpreted either as support for Greene’s position that westernization and high-SES are signs of greater cognitive emphasis, or as support for the idea that an exclusive emphasis on harm in moral judgment is an artifact of westernized cultural approaches. College students may be more “cognitive” than less educated people, but they also tend to be more westernized and to have higher SES than the general population in most countries [12: 625].

the emphasis on harm is connected to the “ethics of autonomy,” which emphasizes not just overall consequences but “harm, rights, and justice.” According to the ethics of autonomy, harm to others and violations of rights are the only types of wrong actions, because apart from those, people have a moral right to control their own lives. So the study is not proposing that subjects who deny that harmless actions are wrong must be relying on consequentialist thinking. Instead, the study assumes that they are relying on a fundamentally deontological, autonomy-based approach that counts harm (to others, not to oneself) as the essential distinguishing feature of the scope of morality. To describe the “not wrong” responses as consequentialist then distorts the position of the study itself. And if “no” responses to the question of whether the action is wrong do not reveal consequentialist thinking, then neither do “yes” responses reveal deontological thinking. Greene himself admits that, “The connection between the tendency to condemn harmless action and deontology is, however, less straightforward and more questionable” than between consequentialism and lack of condemnation [7: 57]. Greene says that the study’s scenario of breaking a promise to one’s dead mother is “downtown deontology” – a violation that most deontological moral theories would condemn – but provides no overall reason to associate moral judgments that harmless actions are wrong with deontological theory. He mentions that “commonsense moralists” as well as adherents of any deontological theory will tend to condemn some harmless wrongs (7: 55), but the arguments I have given above, in the section on problems with Greene’s “rationalization” argument, suggest that it is difficult to strongly correlate any particular deontological moral theory with all and only emotional reactions, and that it may therefore be easier to connect non-philosophical, “commonsense” judgments to these emotions. So Greene fails to show that judgments that harmless actions are wrong depend especially on deontological theory, as well as failing to show that judgments that harmless actions are acceptable depend on consequentialist moral theory.

The failure to draw a strong connection between emotion-based reactions and deontology as a moral theory also undermines one additional argument that Greene offers against deontology. Greene argues that psychological studies show that

People endorse both consequentialist and retributive justifications for punishment in the abstract, but in practice, when faced with more concrete hypothetical choices, people’s motives appear to be predominantly retributivist. Moreover, these retributivist inclinations appear to be emotionally driven. [7: 51]

Even if one grants these claims, the problem is that Greene fails to show that these particular emotion-driven retributive judgments correspond to any retributive *theory* of punishment, let alone any more general deontological moral theory. Using the word “retributive” as a label for both particular impulses to punish and for more carefully developed theories of punishment makes it too easy to assume that the main role of

retributive theories of punishment is to justify, or offer a “rationalization” for, individuals’ desires to harm wrongdoers, in a way analogous to Greene’s hypothesis that deontological theories are rationalizations of emotional alarm reactions to personal harm.

But in much the same way that it was hasty to think that deontological moral theories are rationalizations of emotional reactions to personal harm, it also is hasty to regard retributive theories of punishment as rationalizations of emotional reactions of anger toward wrongdoers. Greene admits that even most retributive theories of punishment allow that “prevention of future harm provides *a* legitimate justification for punishment,” and that the distinguishing feature of non-consequentialist theories is that they maintain that “such pragmatic considerations are not the *only* legitimate reason for punishment, or even the main ones” [7: 50]. If so, then few if any theories can be merely rationalizations of angry, retributive emotions. More importantly, most retributive theories and most actual legal systems serve as much to limit anger-based retribution as to legitimize or endorse it. One main point of attempts to insure impartiality and to follow an established set of legal procedures and sentencing guidelines is to limit the extent to which victims’ or others’ outrage and anger directly dictate the ways in which accused wrongdoers are punished. Even the most famous slogan of retributive legal punishment, “An eye for an eye,” is an attempt to limit punishment to an amount that is proportionate to the crime, rather than a demand for bloody vengeance – that is, if someone puts out your eye, you can demand only that he lose his eye, not that he be killed, tortured, exiled or the like. In general, the harm that a wronged person, or even a witness, wishes to inflict does not necessarily correspond to the punishment that an actual retributive theory deems appropriate. So Greene at the least needs to show a closer connection between feelings of anger and retributive theories of punishment, if he is to prove that the feelings motivate the theories.¹⁴

Although the criticism I have offered of Greene’s attempt to show that retributive (and so deontological) theories of punishment is not as detailed as my criticisms of his other positions, both his argument and my criticism follow a general pattern that is familiar. Even if one grants the empirical evidence that Greene cites, there are philosophical problems in showing that the empirical evidence can really be appropriately directed at the deontological theories that are Greene’s target.

In “The Secret Joke of Kant’s Soul,” Greene relies on several lines of empirically based arguments, and says, “Any one of the results and interpretations described here may be questioned, but the convergent evidence

¹⁴ It is worth noting that although [16] associated damage to the VMPFC with both a lessening of social emotions and with an increase in consequentialist judgments in difficult cases, another study by the same primary investigator [17] showed that patients with the same kind of VMPFC damage acted more vengefully (or emotionally?) in retaliating against unfair offers in the ultimatum game. This raises questions about the exact role of social emotions in both kinds of decisions [21].

assembled here makes a decent case for the association between deontology and emotion...” [7: 59]. But I have suggested that there are serious problems with Greene’s three main lines of argument, and this in turn undermines his picture of several converging bodies of evidence. Whatever the many fruitful results of the recent boom in empirical research on moral judgment, Greene has not yet shown that the results of this research provide reason to favor consequentialist over deontological moral theory.

References

- [1] Bartels, Daniel. 2007. Principled moral sentiment and the flexibility of moral judgment and decision making. *Cognition*. 108: 381-417.
- [2] Gazzaniga, Michael and Jonathan Le Doux. 1978. *The Integrated Mind*. New York: Plenum.
- [3] Greene, Joshua, R. Brian Sommerville, Leigh Nystrom, John Darley, and Jonathan D. Cohen. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*. 293: 2105-2108.
- [4] Greene, Joshua, and Jonathan Haidt. 2002. How (and where) does moral judgment work? *Trends in Cognitive Sciences*. 6 (12): 517-523.
- [5] Greene, Joshua. 2003. From neural ‘is’ to moral ‘ought’: what are the moral implications of neuroscientific moral psychology? *Nature Reviews, Neuroscience*. 4: 847-850.
- [6] Greene, Joshua, Leigh E. Nystrom, Andrew D. Engell, John M. Darley, and Jonathan D. Cohen. 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron*. 44: 389-400.
- [7] Greene, Joshua. 2008. The secret joke of Kant’s soul. In *Moral Psychology, volume 3, the Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. Walter Sinnott-Armstrong, 35-79. Cambridge, Massachusetts, and London: MIT Press.
- [8] Greene, Joshua. 2008. Reply to Mikhail and Timmons. In *Moral Psychology, volume 3, the Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. Walter Sinnott-Armstrong, 105-117. Cambridge, Massachusetts, and London: MIT Press.
- [9] Greene, Joshua, Sylvia A. Morelli, Kelly Lowenberg, Leigh E. Nystrom, and Jonathan D. Cohen. 2008. Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*. 107: 1144-1154.
- [10] Greene, Joshua. 2009. “Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. *Journal of Experimental Social Psychology*, 45: 581-584.
- [11] Haidt, Jonathan. 2001. The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*. 108 (4): 814-834.
- [12] Haidt, Jonathan, Silvia Helena Koller, and Maria G. Dias. 1993. Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*. 65(4): 613-628.
- [13] Kant, Immanuel. 1965. *Critique of pure reason*. Translated by Norman Kemp-Smith from *Kritik der Reinen Vernunft*. New York: St. Martin’s Press.
- [14] Kant, Immanuel. 1996. *The metaphysics of morals.*, Translated by Mary Gregor, from *Die Metaphysik der Sitten*, vol. vi of *Kant’s Gesammelte Schriften*, 203-491. Cambridge: Cambridge University Press.

- [15] Kant, Immanuel. 2002. *Groundwork of the metaphysics of morals*. Translated and edited by Thomas Hill, Jr. and Arnulf Zweig, from *Grundlegung zur Metaphysik der Sitten*, vol. iv of *Kant's Gesammelte Schriften*, 387-463. Oxford: Oxford University Press.
- [16] Koenigs, Michael, Liane Young, Ralphs Adolphs, Daniel Tranel, Fiery Cushman, Marc Hauser, and Antonio Damasio. 2007. Damage to the prefrontal cortex increases utilitarian moral judgments. *Nature*. 446: 908-911.
- [17] Koenigs, Michael, and Daniel Tranel, D. 2007. Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the ultimatum game. *Neuroscience*. 27: 951-956.
- [18] Luce, Mary, James Bettman, and John Payne. 1997. Choice processing in emotionally difficult decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 23: 384-405.
- [19] McGuire, Jonathan, Robyn Langdon, Max Coltheart, and Catriona Mackenzie. 2009. A reanalysis of the personal/impersonal distinction in moral psychology research. *Journal of Experimental Social Psychology*. 45: 577-580.
- [20] Mikhail, John. 2008. Moral cognition and computational theory. In *Moral Psychology, volume 3, the Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. Walter Sinnott-Armstrong, 81-91. Cambridge, Massachusetts, and London: MIT Press.
- [21] Moll, Jorge, and Ricardo de Oliveira-Souza. 2007. Moral judgments, emotions, and the utilitarian brain. *Trends in Cognitive Science*. 11 (8): 319-321.
- [22] Monin, Benoît, David Pizarro, and Jennifer Beer. 2007. Deciding versus reacting: Conceptions of moral judgment in the reason-affect debate. *Review of General Psychology* 11 (2): 99-111.
- [23] Moore, Adam, Brian Clark, and Michael Kane. 2008. Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science*. 19(6): 549-557.
- [24] Nisbett, Richard. and Timothy Wilson. 1977. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*. 84: 231-259.
- [25] Ross, William David. 2003. *The right and the good*. New York: Oxford University Press.
- [26] Schaich Borg, Jana, Catherine Hynes, John Van Horn, Scott Grafton, and Walter Sinnott-Armstrong. 2006. Consequences, action, and intention as factors in moral judgments: An fMRI investigation. *Journal of Cognitive Neuroscience*. 18 (5): 803-817.
- [27] Tetlock, Philip, Randall Peterson, and Jennifer Lerner. 1996. Revising the value pluralism model: Incorporating social content and context postulates. In *Ontario Symposium on Social and Personality Psychology: Values*, eds C. Seligman, J. Olson, and M. Zanna. Hillsdale, NJ: Earlbaum.
- [28] Tetlock, Philip, Orië Kristel, S. Beth Elson, Melanie Green, and Jennifer Lerner. 2000. The psychology of the unthinkable: taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*. 78: 853-870.
- [29] Timmons, Mark. 2008. Toward a sentimentalist deontology. In *Moral Psychology, volume 3, the Neuroscience of Morality: Emotion, Brain Disorders, and Development*, ed. Walter Sinnott-Armstrong, 93-104. Cambridge, Massachusetts, and London: MIT Press.