

THE SECRET JOKE OF GREENE'S FORK

A Thesis submitted to the faculty of
San Francisco State University

AS In partial fulfillment of
36 the requirements for
2017 the Degree

PHIL Master of Arts

.V464 In
Philosophy

by

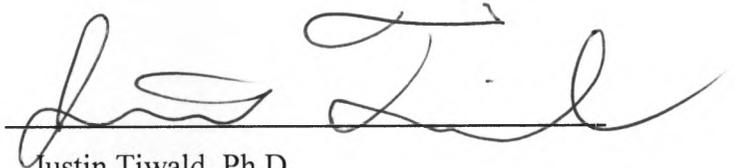
Miguel Angel Venegas

San Francisco, California

May 2017

CERTIFICATION OF APPROVAL

I certify that I have read The Secret Joke of Greene's Fork by Miguel Angel Venegas, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirement for the degree Master of Arts in Philosophy at San Francisco State University.



Justin Tiwald, Ph.D.
Associate Professor of Philosophy



Isabelle Peschard, Ph.D.
Associate Professor of Philosophy



Alan Moore, Ph.D.
Lecturer of Philosophy

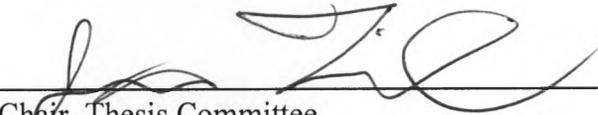
THE SECRET JOKE OF GREENE'S FORK

Miguel Angel Venegas
San Francisco, California
2017

Most of people's everyday judgments are inspired by emotions. Rarely do people rely on their cognitive resources to judge a situation rationally. Can the same be said of people in moral situations? If so, what is the relationship between people's psychological patterns and moral judgments? Contrary to historical stereotypes, research in moral psychology shows deontological judgments are mostly driven by emotions while consequentialist judgments are mostly driven by cognition. Joshua Greene argues that these latter claims are manifestations of two dissociable psychological processes, which is a strong metaphysical claim. If the metaphysical expectation is true, the implications may be philosophically groundbreaking, for it could cast doubt on the normative status of a particular moral theory and, at the same time, elevate the status of another. Before seriously considering the implications, I reexamine the evidence gathered and theoretical commitments used to ambitiously motivate the strong metaphysical claim. In doing so, I undercut Greene's metaphysical claim by arguing for the opposite set of claims, namely, that deontological judgments can be driven by cognition and that consequentialist judgments can be driven by emotions.

Keywords: deontological judgment; consequentialist judgment; emotion;
cognition; moral psychology

I certify that the Abstract is a correct representation of the content of this thesis.



Chair, Thesis Committee

4/27/17

Date

ACKNOWLEDGEMENTS

When first exposed to the wide-ranging and contemporary topics covering the moral psychology literature two semesters ago, I was immediately impressed by the experimental approach to morality. Since, then, I have been slowly consuming some of the main ideas in the field. With the help of my professors, Justin Tiwald, Alan Moore, and Isabelle Peschard, I have turned my inchoate curiosity and fledging ideas into a thesis. Justin, thanks for the insights, encouragement, and editorial suggestions. Alan, thanks for your direction, rigor, and expertise. Isabelle, thanks for your feedback, support, and instruction. Moreover, I would like to thank my friends and family who have not taken offense at my hermit-like behavior these past few months. Lastly, I would like to thank my wife Veronica who has not only encouraged me throughout this project, but who has also allowed me to miss occasional house-hold duties in the interest of letting me write.

TABLE OF CONTENTS

1. Introduction.....	1
2. Literature Review.....	2
3. Defining Moral Judgments.....	7
4. The Trolley Problem.....	8
5. Strong Metaphysical Claim.....	10
6. Theoretical Commitments.....	12
6.1 Dual-Process Theory.....	12
6.1.1 Intuitive Versus Rational.....	13
6.1.2 Emotional Versus Cognitive.....	14
6.2 Sharp Distinctions.....	16
6.2.1 Emotional Distinctions.....	17
6.2.2 Emotional-Cognitive Distinctions.....	18
6.3 Interpreting Greene's Thesis.....	19
7. Explaining the Trolley Problem.....	20
7.1 Hypothesis.....	20
7.2 Neuro-Imaging Evidence.....	21
7.3 Atypical Judgments.....	22
7.4 Reaction-Time Evidence.....	23
8. Method of Criticism.....	25
9. Arguments.....	26
9.1 First Argument.....	26
9.1.1 Evidential Standard.....	27
9.1.2 Missing the Standard.....	27
9.1.3 Empirical Solution.....	28
9.2 Second Argument.....	31
9.2.1 The Crying Baby Dilemma.....	32
9.2.2 A Cognitive Interpretation.....	33
9.2.3 Problems with Greene's Interpretation.....	34
9.2.4 A Theoretically Consistent Interpretation.....	34
9.2.5 Greene's Fork.....	38
9. Objections.....	40
10. Implications/Conclusion.....	44
11. References.....	47

How do people resolve moral situations that force them to choose between two life-threatening options? Do people use reason as their guide before endorsing an option? Or, do people endorse an option because an emotion inspires them to choose it? According to Joshua Greene, the answers to these questions may depend on the kind of judgment being endorsed. In “The Secret Joke of Kant’s Soul,” Greene takes an empirical approach to explain the Trolley Problem’s stereotypical, moral judgments. One kind of moral judgment is characteristic of consequentialism, while the other kind of moral judgment is characteristic of deontology. Greene attempts to show that moral judgments are manifestations of two dissociable psychological patterns. Proving the latter may warrant the claim that he is uncovering the nature of each type of moral judgment. To establish this strong metaphysical claim or [M], he must equally argue for a strong thesis that shows [1] deontological judgments are driven by alarm bell emotions, and [2] consequentialist judgments are driven by cognition and currency-like emotions. Though promising and potentially groundbreaking, I indirectly undermine [M] by showing that some dilemmas fail to meet the same evidential standard cited in the footbridge dilemma. As the consequence of failing to meet the standard, I argue that, by using his own theoretical commitments and (partial) evidence, the opposite set of claims can also be made, namely, that [3] deontological judgments are driven by cognition, and [4] consequentialist judgments are driven by alarm bell emotions. If successful, my arguments will show that Greene’s account does not accurately describe the nature of moral judgments.

Literature Review

In arguing for his thesis, Greene would like to change how we understand consequentialism and deontology. Instead of accepting consequentialism as “more emotional,” he would like to empirically demonstrate that it is more “cognitive.” Moreover, instead of accepting deontology as more “cognitive,” he would like to also demonstrate that it is “more emotional.” If his empirical project is successful, the implications are groundbreaking.

To understand and appreciate Greene’s seminal research, however, it is important to situate him in the moral psychology literature. His research can be viewed as an effort to (partly) resist Jonathan Haidt’s (2001) radical “Social Intuitionist Model” (SIM) and as an effort to adopt Jonathan Baron’s (1994) research on non-consequentialist decision-making. Greene’s research addresses both efforts empirically. In the interest of space, I will only highlight the most notable ideas from each of their research projects.

Haidt’s research flies in the face of dominating rationalist, empirical models. Whereas past rationalistic models claim moral judgments result from deliberation, introspection, reasoning, and/or reflective capacities (Piaget, 1965; Turiel, 1983), Haidt claims most moral judgments result from an entirely different process. Evidence gathered shows that most moral judgments are made quickly, automatically, and unconsciously (Haidt, Bjorklund & Murphy, 2000; Greene & Haidt, 2002). A moral judgment, in this context, is defined as, “evaluations (good versus bad) of the actions or character of a person that are made with respect to a set of virtues by a culture or subculture to be

obligatory” (Haidt, 2001, p. 6). The fact that moral judgments are made quickly, automatically, and unconsciously suggest that people are not relying on reasoning to make decisions. Instead, people seem to be relying on gut reactions or intuitions. An intuition, in this context, is likened to a perceptual process, enabling individuals to evaluate situations, actions, or characters “without any awareness of the mental process leading to the outcome” (p. 6). Haidt’s research on intuitions is thought-provoking because he argues for intuition’s primary role on judgments. “Moral intuitions...,” he tells us, “come first and directly cause moral judgments...” (p. 2). This last point is interesting because, in past research, reasoning has been thought to come first and directly cause judgments. However, in this social intuitionist model, moral reasoning takes on an indirect role, occurring mostly after a moral judgment has already been made. Thus, reasoning is primarily viewed as an “ex-post facto process” and “rarely a direct cause of moral judgment[s]” (p. 3). In most cases, reasoning takes on a subordinate role, formulating reasonable-sounding justifications for already-made judgments or rationalizations. The example given to make sense of both roles is that of a judge and lawyer. Intuition is like a judge, searching for truth and falsehood; reasoning is like a lawyer, trying to build a case (p. 2). Having discussed snippets of Haidt’s research, I will now proceed to explain aspects of Baron’s research, focusing on judgments and the decision-making process.

Baron’s (1994) research on decision-making and judgments stresses the importance of endorsing consequentialist judgments over non-consequentialist

judgments. His argument focuses on the decision-making process. The rationale for this focus is that there is a supposed relationship between decision-making and quality of life. By improving decision-making, people's judgments are advantageously shaped, ultimately helping them achieve and materialize desired goals. To implement this consequentialist standard, Baron first locates the dangers that inhibit it. A common problem found is that there is an unwavering bias for acts of omission than acts of commission, even when the harms are identical. More surprisingly, though, the same bias is repeated when the overall benefit weigh against the act of omission. In other words, a common problem that Baron found is that people often make decisions without considering consequences. Baron suggests that the possible sources for this non-consequentialist tendency may arise from an emotional incitation, an evolutionary impulse, or an employment of the naturalistic fallacy. In any case, Baron argues that, at best, judgments springing from non-consequentialist norms are overused generalizations that, at some point, helped overcome a specific moral situation and, at later point, failed to apply to newer situations. In short, people are trying to solve unprecedented moral problems with outdated, context-specific heuristics. Baron believes it is possible to overcome or mitigate these heuristics' efficacy, but it will require retraining of some sort. Eventually this non-consequentialist dislodgment would allow for a reestablishment of a better set of reasoning norms, which align more readily with consequentialism. Though much more can be said, I believe the discussion on the literature so far will provide insight into Greene's research.

These two discussed research projects provide a template for Greene's empirical framework. With Haidt, we learn that most judgments are based on emotional intuitions (Greene & Haidt, 2002). Rarely are judgments made directly from reason. However, unlike Haidt's emphasis on intuitions and de-emphasis on reason, Greene's research centers on cognition (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Paxton, Ungar & Greene, 2012). His slight emphasis on cognition is defensible when we consider that SIM relies on a dual-process theory (hereafter, DPT). (I will discuss DPT in greater detail in subsequent sections). In a broader sense, then, Greene is assessing the scope of SIM by investigating the relevant role of cognition when making judgments. A requirement for achieving such empirical project requires, first, differentiating among various kinds of moral judgment. Greene (2008), in fact, complains that,

Haidt does not distinguish among the various approaches to ethics familiar to moral philosophers: consequentialism, deontology, virtue ethics, etc. Rather, his radical thesis is intended, if only implicitly, to apply equally to the adherents of all moral philosophies, though not necessarily well to moral philosophers as a group. (p. 36)

To paraphrase, Greene thinks Haidt's thesis is broad and unspecific, applying equally to all philosophies and moral judgments. Eager to spell out the stark differences among moral philosophers, Greene adopts Baron's classification of moral judgments. Since Baron makes a psychological distinction between two types of judgments (consequentialist and non-consequentialist judgments), this made differentiation allows

Greene to explore the relationship between moral judgments and evoked psychological processes. In the coming section, Greene will argue deontological judgments are driven by emotion while consequentialist judgments are driven by cognition. Before explaining how all these claims unravel, let's consider how Greene (2008) differs from some historical stereotypes. He tells us,

There is the historical stereotype, according to which consequentialism is more emotional (emerging from the “sentimentalist” tradition of David Hume (1740/1978) and Adam Smith (1759/1976)) while deontology is more “cognitive” [encompassing the Kantian “rationalist” tradition (Kant, 1959)]. (p. 41)

In the sentimentalist camp, he includes philosophers such as David Hume and Adam Smith and, on rationalist camp, he specifically has in mind Immanuel Kant. Historically, Immanuel Kant's moral philosophy may have been stereotyped as the paragon of rationality or abstract reasoning. Very few, if any, would have imagined that his entire deontological framework could be sourced from a psychologically, contingent feature. However, Greene wants to convince us that Kant entirely misunderstood the very nature of his own philosophical inspiration. Though Kant may have thought his moral philosophy was a direct product of reason, Greene thinks it is a direct product of a specific kind of emotion (i.e., alarm bell emotion). If it turns out that, in general, the deontological framework is a product of alarm bell emotions, then not only will it alter our esteem for deontological theorizing, but also, in some degree, alter our confidence in

solving moral problems using that approach. One may lose esteem for this deontological approach because its moral conclusions would be nothing more than mere rationalizations. That is, any reasons, justifications, or arguments offered to solve moral problems would be secondary, incidental, and irrelevant. Arguably, at that point, deontology would lose its credentials as “a school of normative thought” (Greene, 2008, p.36; cf. Greene & Cohen, 2004). Seriously dwelling and, perhaps, accepting these implications should be postponed until one thoroughly examines Greene’s evidence and theoretical commitments.

Defining Moral Judgments

In this section I will introduce stereotypical judgments and, then, the dilemmas that inspired Greene’s empirical project. Greene attempts to explain two stereotypical, moral judgments that occur when people contemplate, what has been called, the Trolley Problem (Foot, 1967; Thomson, 1986). Explaining the stereotypical patterns are an important part of Greene’s empirical project because they show the influencing roles of emotion and cognition on judgments. One stereotypical pattern is characteristic of deontological moral theories while the other is characteristic of consequentialist moral theories. Greene (2008) defines moral judgments functionally. A deontological judgment, Greene explains, aligns itself more readily with deontological theories whose sole emphasis is respect for rule-based, moral concepts such as individual rights and humanitarian duties. Moral phrases such as “the ends don’t justify the means,” “duty for duty’s sake,” and “it’s wrong despite the benefits” are accurate, if not succinct,

characterizations of this deontological, normative theory. In contrast, Greene thinks consequentialist judgments occur when considering the overall consequences of an action. Though still acknowledging the rights of individuals and duties of humanity, a consequentialist moral theory is distinguishable from the deontological moral theory in that it is concerned with the welfare of all parties involved. At times, the interests of the many will override individual interests. As the result, consequentialism is more flexible than the rigid imperatives imposed by the former theory. Moral phrases such as “the ends do justify the means” and “better to save more lives at the expense of one” are phrases that more readily align with consequentialist normative theories. From here on whenever subjects choose behavioral responses in the trolley dilemmas that typify these moral characterizations, Greene will count them as functional representations of the moral theory itself.

The Trolley Problem

At this point, I will introduce the two well-known dilemmas and the problem they pose. I will begin with the switch dilemma, then the footbridge dilemma.

The Switch Dilemma

A runaway trolley is headed for five people who will be killed if it proceeds on its present course. The only way to save these people is to hit a switch that will turn the trolley onto a side track, where it will run over and kill one person instead of five. Is it okay to turn the trolley in order to save five people at the expense of one? (Greene, 2008, pp. 41-42)

The Footbridge Dilemma

A runaway trolley threatens to kill five people, but this time you are standing next to a large stranger on a footbridge spanning the tracks, in between the oncoming trolley and the five people. The only way to save the five people is to push this stranger off the bridge and onto the tracks below. He will die as a result, but his body will stop the trolley from reaching the others. Is it okay to save the five people by pushing this stranger to his Death? (p. 42)

In these trolley problems, saying “yes” to the first dilemma and “no” to the second dilemma aligns one with people’s stereotypical responses. Most subjects (63%), that is, typically endorse the option that requires pulling the switch in the switch dilemma; and, likewise, most subjects (69%) typically endorse the option that requires refraining from pushing the large man off the bridge in the footbridge dilemma (Tiberius, 2014, p. 193). More importantly, as a consequence of these findings, we learn that each dilemma inspires different kinds of moral judgments. For instance, since the yes response or pulling the switch results in the saving of five lives at the expense of one, it is characteristic of consequentialism. Moreover, since the no response or not pushing the large man off the bridge results in the loss of five lives at the deliverance of one, it is characteristic of deontology (it’s wrong despite the benefits!). These stereotypical responses are interesting and problematic because both dilemmas are nearly identical, yet they generate different moral judgments. Though both dilemmas pressure subjects to save five lives at the expense of sacrificing one, the acceptability of potentially harming a

person to achieve such altruistic aim drastically changes. Why do these dilemmas yield different moral judgments? Greene believes emotion and cognition are the causal factors responsible for producing these stereotypical judgments.

Strong Metaphysical Claim

In offering a solution to the Trolley Problem, Greene aims to convince us of a strong metaphysical claim, namely, that each moral judgment is a manifestation of a psychological natural kind. We get a glimpse of this reductive aim when first introducing his thesis:

I will argue that deontological judgments tend to be driven by emotional responses, and that deontological philosophy, rather than being grounded in moral reasoning, is to a large extent an exercise in moral *rationalization*. This contrast with consequentialism, which, I will argue, arises from rather different psychological processes, ones that are more ‘cognitive’ and more likely to involve genuine moral reasoning. (Greene, 2008, p. 36)

In this passage, Greene’s interest in exploring the causes of judgments is to confer on each judgment a specific reasoning status (genuine reasoning versus disingenuous reasoning, i.e., rationalizations). Deontology and its respective judgments are products of emotions. So, if reasoning does occur in this kind of philosophical thinking, it is nothing more than exercise of “rationalizations” or an attempt to justify gut reactions (i.e., intuitions). Reason, of this kind, is retroactive. In contrast, consequentialist thinking involves “genuine reasoning” in so far as subjects rely on their cognitive processes to

judge moral situations *in vivo*. This real-time reasoning occurs alongside of and at the expense of emotional intuitions. Reason, of this other kind, is proactive.

Greene's primary interest in assigning judgments a specific reasoning status is to say something about the nature of judgments. This nature-assigning motive is prevalent in various areas of his article (Greene, 2008, p. 38; pp.39-40; p. 63). In the interest of space, however, I will focus on one passage where this motive is stated explicitly. Greene tell us,

I believe it is possible that philosophers do not necessarily know what consequentialism and deontology really are. How could this be? The answer, I propose, is that the terms 'deontology' and 'consequentialism' refer to *psychological kinds*. I believe that consequentialist and deontological views of philosophy are not so much philosophical inventions as they are philosophical manifestations of two dissociable psychological patterns, two different ways of moral thinking... (p. 37)

In this passage, I believe, Green shows concern for the stronger metaphysical claim: moral philosophies, as well as their respective judgments, are concrete manifestations of "dissociable psychological patterns" or "two different ways of moral thinking." They are, in short, psychological natural kinds [M]. Before making sense of Greene's evidence and its relationship to [M], I must expound on his theoretical commitments.

Theoretical Commitments

Dual Process Theory

Greene is committed to a theory which claims there are two psychological processes informing our decision making and influencing our judgments and actions (Cushman, Young & Greene, 2010). The processes are distinct, each possessing their own individuating characteristics and functions. Often the processes work in tandem. Other times, however, these processes noticeably conflict; hence asserting their independence and highlighting a tension between two different ways of solving problems (Greene & Haidt, 2002).

The discussion on DPT is messy, however. It is composed of loosely defined and seemingly disintegrated components (i.e., intuition, rationality, cognition, and emotion). Despite its apparent complexity, Greene has conducted various studies that lends modest support for DPT's establishment (Greene, 2007; Greene et al., 2008), though many difficulties remain (Cushman, Young & Greene, 2010). One of the main difficulties with this theory is integrating its various components in a cohesive, understandable framework (Cushman, 2013; Cushman, Young & Greene, 2010). Another difficulty, however, concerns the general distinctions, such as the supposed distinctions between cognition and emotion or between what is intuitive and rational (Kahneman, 2011). In this paper, I will not focus on these difficulties since it would unnecessarily lead us away from discussion at hand. Instead, I will focus on the distinguishing marks of DPT. I will explain them one-by-one, contrasting them with their opposites.

Intuitive versus Rational. The phenomenon of “moral dumbfoundedness” has often be used as evidence to show intuition is separate from rational thinking (Björklund, Haidt & Murphy, 2000). Moral dumbfoundedness is “the stubborn and puzzled maintenance of a judgment without supporting reasons” (p. 1). With the intent of instigating this phenomenon, Haidt et al. (2000) conducted a study to test people’s approval or disapproval of moral and non-moral situations, such as scenarios involving cannibalism, incest, and roaches. Immediately after presenting these and other scenarios, subjects were asked the following evaluative questions: “Was it wrong for Jennifer to eat human flesh?” “Was it wrong for Julia and Mark (siblings) to have sex?” and “Would you like to drink juice/water after a sterilized cockroach has been dipped into it?” Many subjects had robust, negative responses (or evaluations) to these posed questions, but lacked the ability to justify their immediate gut reactions. In some cases, subjects avowed that “harm” was the reason for not accepting the act in question. But such post-hoc reasoning is questionable because the scenarios were “engineered...to make it extremely difficult to find strong arguments to justify [those] intuitions” (p.5). Because the scenarios were designed to control for harm-based reasons, many subjects were left dumbfounded, saying things like, “I know it’s wrong, but I can’t come up with a reason why” (p. 10). Being dumbfounded shows that, in some cases, the immediate, automatic, quick, process makes evaluations without the assistance of reason. So, when speaking of intuition, I take it to mean something like this unreflective, effortless, quick, unconscious process.

In contrast to this intuitive process, there is a rational process that is “slow, effortful, and voluntary” (Cushman, Young & Greene, 2010, p. 12). Though partly agreeing with Haidt’s view on rationality, Greene distinguishes himself by emphasizing to a greater degree its direct and positive role on our judgments. Greene likens the rational process to a controlled process, such as experienced by subjects contemplating the Stroop Task (Macleod, 1991; Stroop, 1935). In this study, subjects were presented with an array of words, naming various colors (i.e., red, blue, green, etc.). Each word, however, was typed in different ink colors. Subjects were then given the task to name the ink color. Though the task sounds simple enough, the study showed that it was very difficult because of the temptation to read the word rather than the ink color (Greene, Nystrom, Engell, Darley & Cohen, 2004). To complete the task successfully, subjects had to deliberately and consciously focus on the goal, using a controlled process. Like in this study, Greene et al. (2004) believe that the rational process is an exercise of mental discipline, requiring an uncanny ability to “guide attention, thought, and action in accordance with goals or intentions, particularly in the face of competing behavioral pressures (Cohen et al. 1990; Posner & Snyder 1975; Shiffrin & Schneider 1977)” (p. 390). Given this positive view of rationality, Greene thinks that, in some situations, some can overcome the more immediate, gut reactions by channeling their rational processes.

Emotional versus Cognitive. Greene distinguishes emotion and cognition with respect to their behavioral representations. A species of emotions, also known as alarm bell emotions, represent inflexible behaviors. They are like emotional intuitions because

they automatically presuppose “a particular behavioral response” (Cushman, Young & Greene, 2010, p. 12). There is an evolutionary story that explains why certain actions are considered inflexible or powerfully motivating, but I will only mention that these alarm bell emotions are like Kantian imperatives—unyielding, uncompromising, and rigid. When a personal factor evokes these alarm bell emotions, people register those emotional responses as “No! Don’t do that!”

In contrast to emotion’s inflexible behavior representations, cognitive behavioral representations are “inherently neutral.” They do not automatically trigger responses or dispositions. This capacity is important because it helps people do every day things like crossing the street or taking a taxi. Most of the time, Greene tell us, people need to approach (i.e., like when you need a taxi) and avoid cars (like when a car is speeding out of control and headed your way). This dual capacity (to avoid or approach cars) would not be possible if cars were not represented in a behaviorally neutral way (p. 40). There are other features that distinguish cognition from stubborn emotions. Cognition allows subjects to engage in time-consuming activities such as “cost-benefit-analysis” and “abstract reasoning.” In fact, Greene explicitly states, “the ‘cognitive’ process is especially important for reasoning, planning, manipulating information in working memory, controlling impulses, and ‘higher executive functions...’” (p. 40). The cognitive process takes into consideration all, if not most, aspects of a situation. It can “mix around” and “recombine” behavioral representations as the situation requires.

Having explained DPT's four critical components separately may give the impression that all components function independently from each other. These components hold interconnected relationships. Greene assumes that intuition and emotion (hereafter, System 1) are conjoined, and that rationality and cognition are also conjoined (hereafter, System 2). Since my aim is to understand Greene's research as best as possible, I will assume that this way of conjoining each component is the correct interpretation.

Sharp Distinction

Greene makes sharp distinctions between emotion and cognition, and even makes distinctions between emotions themselves. These distinctions are crucial because they are presupposed by his metaphysical claim. Though he does make sharp distinctions, in some portions of his article Greene (2008) denies their existence. In fact, he explicitly states in the beginning of his article: "I don't believe that either approach is strictly emotional or 'cognitive' (or even that there is a sharp distinction between 'cognition' and emotion)" (p. 42). He claims, more specifically, that he is influenced by the Humean portrayal of moral judgments, according to which, "...all moral judgments (including consequentialist judgment) must have some emotional component" (p. 42). If there are no distinctions, then there are no purely cognitive and purely emotional judgments.

Emotional Distinctions. Despite this explicit denial of a sharp distinction, in other portions of his article, Greene speaks as if there are distinguishing marks—be it real, provisional, or quasi. For instance, he affirms that different kinds of emotions are intrinsic to each moral judgment: “I suspect that the kind of emotion that is essential to consequentialism is fundamentally different from the kind that is essential to deontology, the former functioning more like currency and the latter functioning more like an alarm” (p. 41). Alarm bell emotions are essential to deontological judgments; currency-like emotions are essential to consequentialist judgments. Moreover, these two emotions, he tells us, are “fundamentally different.” Unlike alarm bell emotions, which are depicted as motivationally unyielding, currency-like emotions are “designed to add a limited measure of motivational weight to a behavioral alternative...” (Cushman, Young & Greene, 2010, p. 12). When deciding between two competing options, currency-like emotions alter the momentum by either adding or subtracting motivational points. In other words, they are flexible and, for that reason, negotiable, unauthoritative, and tractable. Further, if disobeyed or disregarded, these emotions do not “leave one with a sense of discomfort” (p. 12). Here’s an example of how currency-like emotions may operate in one’s daily life. Imagine deciding between getting ice cream on a hot summer day (which gives one reason to pursue the ice-cream truck) or maintaining a slim, model-looking body (which gives one reason to lounge by the poolside). Choosing either is negotiable, for there are no “resolute commands” pulling one to endorse either option (Cushman, Young & Greene, 2010).

Emotional-Cognitive Distinctions. Another way Greene makes distinctions is when discussing cognition and emotion, focusing primarily on their unique behavioral representations. As expressed in DPT section, cognition deals with flexible behavioral representations while emotion deals with inflexible behavioral representations. The need to represent flexible behaviors makes sense when considering cognition surveys all aspects of a situation before deciding what action to endorse. Given the discussion on currency-like emotions, we can suppose that these emotions are suitable for this cognitive process, since they allow cognition to fluidly examine the possibilities without being compelled to endorse a resolute action. Unlike the cognitive process, alarm bell emotions represent behaviors inflexibly, requiring subjects to endorse an action quickly, automatically, and unconsciously. As mentioned, these alarm bell emotions issue commands, “[driving] the agent unequivocally toward a particular behavioral response” (Cushman, Young & Greene, 2010, p. 12).

Given that clear-cut distinctions are drawn between judgments, it is difficult to interpret Greene’s explicit denial. I assume, hereon, Greene is tacitly committed to sharp distinctions. Besides, without them, I don’t see why his thesis would be historically interesting and empirically groundbreaking. Greene, in general, is concerned with solidifying the nature of each moral judgment, as he believes judgments are “dissociable psychological manifestations” or “psychological natural kinds.” It is difficult to picture how this reductive aim should unfold if there are really no such distinctions. Without pressing any further on this issue, I assume the existence of sharp distinctions.

Interpreting Greene's Thesis

The discussion on DPT indicates that systems are independent and, at times, competitive. The discussion on the sharp distinction indicates that components, such as emotion and cognition, can be either compatible or incompatible. Given these theoretical commitments and the desire to motivate [M], an interpretive problem with his thesis emerges. Is he arguing for a strong or weak thesis? His weak thesis states that deontological judgments are mostly (or tend) driven by emotion and consequentialist judgments are mostly (or tend) driven by cognition (a weaker thesis). His weaker thesis is problematic for a couple of reasons. The first problem concerns his sharp distinctions. Since Greene has assigned each judgment with a specific emotion, his thesis must mention the role of emotion for each judgment. His weaker claims (or weaker thesis) fail to mention the emotional distinctions because they don't specify which emotions he is referring to, thus, giving the impression that he is only referring to one and giving the impression that there are no such distinctions (i.e., the same emotion can drive both judgments). The second problem concerns [M]. Since Greene has made a strong metaphysical claim, the proactive reasoning involved when making the consequentialist judgment cannot be the same reasoning involved when making the deontological judgment. The same can be said for emotion. The emotion driving deontological judgments cannot be the same emotion driving consequentialist judgments. For, in that case, Greene cannot diagnose the state of Kant's soul nor can he place consequentialism

on a pedestal. Thus, Greene's thesis requires stronger and more specific claims (namely, claims [1] and [2]).

Fortunately, Greene makes sharp distinctions that commit him to a stronger thesis. Deontological judgments and consequentialist judgments, he tells us, are each driven by distinct emotions. Moreover, the emotions are either compatible or incompatible with cognition. The currency-like emotion that partly drives consequentialist judgments is compatible with cognition. For cognition is most compatible with an emotion that is also flexible and unauthoritative. In contrast, the alarm bell emotion that drives deontological judgments is incompatible with cognition because that emotion is inflexible, authoritative, and powerful, which precludes cognition from performing its executive duties. Given these distinctions, I believe, Greene does have the ingredients to argue for a stronger thesis: [1] deontological judgments are driven by alarm bell emotions and [2] consequentialist judgments are driven by cognition and currency-like emotions. From here on, I will assume Greene is committed to this stronger thesis.

Explaining the Trolley Problem

Hypothesis

So how does Greene resolve the trolley problem? Greene (2008) hypothesizes that each dilemma evokes a predominant psychological response because of the presence or absence of a personal factor. In the footbridge dilemma, subjects are presented with an option to push a stranger to save five lives. To follow through with such option, subjects must be okay with the "up close and personal" nature of that action. The thought of even

entertaining such personal action, Greene tells us, evokes an alarm bell response (No! Don't push him!). In the absence of such personal factor, like in the switch dilemma, the alternative, cognitive psychological process is evoked. He calls this alternative dilemma an impersonal dilemma. There is an intuitive appeal for calling it impersonal or not personal. Instead of having the face-to-face interaction, the impersonal dilemma requires subjects to bring about the death of a stranger in an indirect and remote way, i.e., by pulling a switch versus by using hands to push. Without dwelling too much on what quality (or qualities) distinguishes the personal from the impersonal factor, I should mention that Greene thinks this personal factor incites a powerful and primitive emotional response. In fact, he even posits that, "...this difference in emotional responses explains why people respond so differently to these two cases" (p. 43). With that, I will now proceed to explain how his evidence may confirm this last point.

Neuro-Imaging Evidence

When comparing neuro-images, Greene (2008) finds different psychological activities in subjects who contemplate "personal moral dilemmas" and "impersonal moral dilemmas." Subjects contemplating personal dilemmas, such as in the footbridge dilemma, show an increase of psychological activity in regions usually associated with emotion such as the posterior cingulate cortex, the medial prefrontal cortex, and the amygdala. In contrast, subjects contemplating impersonal dilemmas, such as switch dilemma, demonstrate an increase of psychological activity in regions "classically" associated with cognition such as dorsolateral prefrontal cortex and inferior parietal lobe

(p. 44). As the result of these varied psychological manifestations found in subjects, Greene believes his hypothesis can help explain why both dilemmas yield different stereotypical judgments. Since subjects contemplating the footbridge dilemma yields mostly deontological judgments, and since subjects contemplating the switch dilemma yields mostly consequentialist judgments, these different stereotypical judgments may be caused by subjects' discrete psychological processes. That is, given that neuro-imaging data demonstrates emotion is pervading subjects in the footbridge dilemma, it seems reasonable to attribute emotion as the cause of those made deontological judgments. Further, given that cognition is highly active in subjects who are contemplating the switch dilemma, it also seems reasonable to attribute cognition as the cause of those made consequentialist judgments.

Atypical Judgments

These general associations, however, overlook a minor yet important detail. Not all subjects conform to the hypothesis. For example, in the footbridge or personal dilemma, Greene (2008) finds that a few subjects endorsed the consequentialist judgment or the atypical judgment. That is, they endorsed the action that amounted to pushing the stranger off the bridge. This is strange because his hypothesis suggests that in the presence of the personal factor and the emotionally intense activity, we should expect subjects to endorse the deontological judgment. Since his hypothesis is, in effect, making a prediction of what we should observe under specified conditions, there is an implied

causal story. Thus, when subjects do meet the specified conditions but not the prediction, not only is the causal story undermined but so is his hypothesis.

Anticipating my complaint, Greene offers support that suggests there are cross-correlations. Thus, he provides us with an explanation, or alternative causal story, that, in my opinion, wrestles with the atypical judgments. His explanation relies on an overriding effect and reaction time evidence. For instance, when explaining why nonconforming subjects in the footbridge dilemma failed to endorse the stereotypical, deontological judgment, Greene proposes that cognition overrides the powerful emotion that inclines subjects to choose otherwise. If, in fact, this overriding effect is occurring, he continues, then we should expect speed differences between made judgments. The expectation for this reaction time evidence (hereafter, RT) primarily stems from Greene's commitment to dual-process theory (hereafter DPT). Note that neuro-images alone are insufficient to establish [M] because it only supports his weaker thesis, namely, that deontological judgments are mostly driven by emotion, and that consequentialist judgments are mostly driven by cognition. Greene needs to relate this set of data (neuro-images) with another set of data (RT evidence) for his stronger thesis. Whenever Greene relates these two, in my opinion, [M] will be empirically established.

Reaction-Time Evidence

Relating two set of data is important because there are some subjects in the footbridge dilemma who did not conform to the stereotypical pattern and, instead, endorsed the atypical judgment. This is problematic since Greene wants to establish [M].

What, then, explains subjects, who meet the above criteria, yet don't endorse the deontological judgment? Greene, thinks cognition is responsible for the few subjects who endorsed the atypical judgment. He, thus, interprets the atypical judgment as a sign of cognition's capacity to "override" emotions:

People tend to have emotional responses to personal moral violations, responses that incline them to judge against those actions...someone who judges a personal moral violation to be appropriate...will most likely have to override an emotional response in order to do it. (Greene, 2008, p. 44)

Under emotion's reign, cognition must exert itself in such a way that it dominates emotion. Evidence for cognition's dominating effort was found by comparing the RTs of subjects' made judgments. Subjects who endorsed the consequentialist judgment, in the footbridge dilemma, took longer than subjects who endorsed the deontological judgment. The discussion on DPT offers a handy interpretation of why the RT differences are empirically relevant. System 1 was characterized as a process that quickly, automatically, and unconsciously produces judgments. Since deontological judgments are allegedly driven by this system, we should expect these judgments to occur quicker than consequentialist judgments. Further, System 2 was characterized as a controlled process that, among many other things, enables subjects to consciously guide attention to execute desirable goals. Given its capacities, System 2 is slower at solving problems. Since consequentialist judgments are allegedly driven by this system, we should expect these judgments to take longer. Thus, showing that deontological judgments were made

quicker than consequentialist judgments supports [M], for each moral judgment is driven by discrete psychological processes. Further, it supports his assumption that psychological processes compete in such a way that one process overrides the other (i.e., the overriding effect).

Method of Criticism

Establishing [M] requires good evidence. Per Greene's standards, good evidence includes a combination of (a) neuro-images and (b) reaction times. The first piece of evidence associates either emotion or cognition to a specific stereotypical judgment. The second piece of evidence associates the remaining process to a specific atypical judgment. If only (a) is offered as evidence, then Greene could only substantiate his weaker thesis because he won't be able to establish associations without, at the same time, excluding the influence of the alternative process. His weaker thesis, as I said before, does not distinguish between emotions and thus allows one to argue that the same emotion can motivate both judgments (hence, ignoring his sharp distinctions). Consequently, if Greene can't distinguish emotions, then his weaker thesis will also fail to distinguish proactive from retroactive reasoning, thus, allowing one to argue that all judgments can be caused by either proactive or retroactive reasoning (which makes him unable to reverse historical stereotypes). However, if Greene does provide RT evidence, then we have good reason to believe that the speed differences do provide support for his stronger thesis and [M]. Given (a) and (b), one could say that the footbridge dilemma passes the test for good evidence, for Greene offers two sets of data for this dilemma.

Thus, I will compare the evidence from other dilemmas to the same exemplary evidence found in the footbridge dilemma. Whenever the evidence for other dilemmas fail to meet the standard established in the footbridge dilemma, I will show how missing the standard compromises Greene's strong thesis and problematizes his theoretical commitments.

Further, I will also show how the lack of evidence, in one way or another, allows one to argue for the opposite set of claims [3] and [4]. Again, we should expect someone to be able to argue for the opposite set of claims because establishing associations (between judgments and evoked psychological processes) requires that one can, at the same time, exclude the alternative process from intermingling. If I am correct about my arguments and interpretation, it will show that Greene is not getting at the nature of judgments.

Arguments

First Argument

In this section I argue that deontological judgments can be driven by cognition. To argue for this, I will first show that one cannot apply the same overriding-effect rationale to explain the occurrence of atypical judgments in the switch dilemma. For, in this dilemma, there is no RT evidence. Even loosening the evidential requirement is unhelpful. For, if the overriding effect is occurring, it is inconsistent with Greene's hypothesis. As the result, we are left without an explanation for the atypical judgments. Offering an explanation, however, allows one to attribute the occurrence of the atypical

judgment to the dominant psychological process, which also drives the stereotypical judgment.

Evidential Standard. In the footbridge dilemma positing the occurrence of an overriding effect was important because it explained why consequentialist judgments took longer to make than emotionally driven deontological judgments. Though the overriding effect neatly explains the significant speed differences between made judgments in that dilemma, perhaps it doesn't always adequately explain why other nonconforming subjects, in a different dilemma, would endorse the atypical judgment.

Missing the Standard. Take the switch dilemma as an example. If we try to apply the same overriding-effect rationale to explain the occurrence of atypical judgments, we find that there is no RT evidence to support such application. In other words, because there are no speed differences between made judgments, it is presumptuous to think the atypical judgment is explainable by the same overriding effect. Much to my surprise, Greene (2008) seems unalarmed by the fact that there are no speed differences. On the contrary, he reassures us that we should not expect the same speed distinction as in the personal dilemma since this impersonal dilemma does not possess an emotional response:

At the same time, we have no reason to predict a difference in reaction time between "yes" and "no" answers in response to impersonal moral dilemmas like the trolley case because there is, according to this model, no emotional response (or much less of one) to override in such cases. (p. 44)

Part of the reason why Greene thinks speed indifferences are not an issue for this switch dilemma (or, what he calls, the trolley case) is because this impersonal dilemma lacks the emotionally-conducive personal factor. He is obviously referring to his hypothesis here. When there is a personal factor, we expect the relevant, alarm bell emotion to be activated. The absence of emotion, however, is troublesome. Given that there were a handful of subjects who endorsed the atypical, deontological judgment, we are left without an explanation for why some subjects endorsed the atypical judgment. If emotion is not responsible for the atypical judgments, then, what is?

Empirical Solution. I believe I can offer an empirically plausible explanation for this anomaly. I argue cognition is responsible for subjects' endorsement of the atypical, deontological judgment. There are a few reasons. For now, I will stress the neuroimaging evidence. Already Greene has told us that all subjects who contemplated the impersonal or the switch dilemma showed unusual cognitive activity. This cognitive activity in subjects, we are told, is evoked by the impersonal character of the dilemma. That is, when the evolutionary-stimulant (or personal factor) is absent, the alarm bell emotion fails to be triggered, and cognition is permitted to reign, free from obstructions. Thus, there is no alarm bell emotion motivating any of the subjects. Therefore, both moral judgments must be driven by cognition.

Admittedly Greene is unclear about how to properly conceptualize subjects' emotional responses in the impersonal dilemma. In the above passage, he says, all subjects contemplating the dilemma in question experience "no emotional response"; but

parenthetically he says, subjects may experience “much less of [an emotional response].” I have shown that it is problematic to assume the former because it leaves us without an explanation for the atypical judgments. I imagine Greene could argue for the second interpretation of emotion. Perhaps it is an exaggeration to say subjects have no emotional response, but rather that they experience much less of an emotional response. If that is the case, it would, perhaps, explain why only a small portion of subjects endorsed the deontological judgments.

Presumably the emotion Greene has in mind, here, is the special alarm bell emotion that is stimulated by personal dilemmas. His hypothesis, however, prevents him from playing the alarm-bell card. Impersonal dilemmas fail to stimulate the alarm bell emotion; thus, if there is any emotion involved in this dilemma it must be a weaker emotion because it lacks the personal factor. Further, given what we know about emotional distinctions, weaker emotions more accurately describe currency-like emotions. Thus, I take a parsimonious interpretation. It could be that the active emotion is, in fact, a currency-like emotion or the Humean portrayal of emotion. If I am correct in my parsimonious interpretation, then the emotional responses in question are irrelevant since they may only apply to consequentialist judgments. Thus, by accepting that all subjects experience a much less emotional response, it does not follow that it is the alarm bell emotion, which would only motivate subjects to endorse the deontological judgments. But maybe this is the wrong way to interpret this dilemma. As a possible objection, Greene might say there are two kinds of emotions in this dilemma, one that

motivates the consequentialist judgment (i.e., currency-like emotions) and another that motivates the deontological judgment (i.e., alarm bell emotions).

Responding this way, once again, takes us back to Greene's overriding effect, which requires RT evidence. Yet, as mentioned, there is no evidence of this kind (Greene, 2008). Let's entertain the overriding idea for second, however. Suppose that two emotions are involved across subjects, what should we expect given Greene's already mentioned theoretical commitments? For one, we should expect these emotions to differ in their behavioral representations. Those subjects who experience an alarm bell emotion should experience an inflexible, behavioral representation; while those who experience a currency-like emotion should experience a flexible, behavioral representation, which may invite cognition to cooperate with it. Second, if some subjects contemplating the impersonal dilemma are, indeed, singularly motivated by this alarm bell emotion while others are singularly motivated by the currency-like emotion, then we should have some empirical evidence to support these emotional distinctions. Greene does not provide us with neuro-images to show us that subjects who endorsed deontological judgments manifested higher levels of emotional activity in comparison to subjects who endorsed the consequentialist judgment. If there were neuro-images to show differences in emotional activity, then perhaps we would have reason to accept that the emotion driving the consequentialist judgment is not the same one driving the deontological judgment. Since there is no evidence to show us that there are emotional distinctions, one could argue that both judgments are driven by the same currency-like emotion and not an

alarm-bell emotion. For, if only deontological judgments were driven by alarm bell emotions, then there would be RT evidence to substantiate that latter claim. The absence of evidence gives us no reason to believe that an overriding effect, like one discussed in the personal dilemma, is at work in this impersonal dilemma, for there is no alarm bell emotion to override. Hence, we are still left with an empirical anomaly. The only way to explain it, I argued, is by supposing atypical judgments (i.e., deontological judgments) were made under the influence of cognition as well.

Second Argument

In this section, I argue that consequentialist judgments can be driven by alarm bell emotions (per Greene's evidence and theoretical commitments). The dilemma I will use to demonstrate this latter claim is the crying baby dilemma. I argue for this claim by, first, demonstrating why Greene's interpretation of the dilemma is problematic. His interpretation, I mention, leaves us wondering where "cognition" gains its out-of-the-blue motivating force. Further, since his interpretation ignores the role of emotion, it also leaves us wondering what, if any, emotions are involved in this dilemma and how they relate to cognition. Given this set of problems and, not to mention the lack of evidence to support [M], I attempt to provide a much more consistent interpretation given his theoretical commitments. Unfortunately, my attempt to maintain a consistent picture of consequential judgments exposes another problem. As it turns out, currency-like emotions may be undistinguishable from the alarm bell emotions. Thus, this problem

opens the possibility of arguing that consequentialist judgments can be driven by alarm bell emotions.

The Crying Baby Dilemma. Unlike other dilemmas, this personal dilemma is labeled “difficult.” Acknowledging its empirical results demonstrates its inherent difficulty. For instance, everyone takes a long time to endorse any judgment. This means that each option is just as attractive or appealing. Two, this dilemma yields no stereotypical judgment (Cushman, Young & Greene, 2010). It is just as likely for subjects to endorse either the deontological judgment or the consequentialist judgment. What makes this dilemma so difficult? See for yourself: one option requires subjects to smother a baby in the interest of saving an entire village, the obvious consequentialist choice (hereafter, village-saving option); the other option risks the lives of an entire village for the temporary deliverance of the infant’s life, the obvious deontological choice (hereafter, baby-saving option). Let’s get a firsthand acquaintance with the crying baby dilemma:

It is wartime, and you and some of your fellow villagers are hiding from enemy soldiers in a basement. Your baby starts to cry, and you cover your baby’s mouth to block the sound. If you remove your hand, your baby will cry loudly, the soldiers will hear, and they will find you and the others and kill everyone they find, including you and your baby. If you do not remove your hand, your baby will smother to death. Is it okay to smother your baby to death in order to save yourself and the other villagers? (Greene, 2008, p. 44)

Ultimately what Greene wants to show is that two independent systems are in competition. To convince us of this, he must first offer us an interpretation of this dilemma.

A Cognitive Interpretation. As we would expect from someone who is trying to establish [M], Greene (2008) sees this dilemma as a competition between an alarm bell emotion and cognition. He thinks that subjects are torn between a “prepotent, negative emotional response” (or an emotion that says “No! Don’t kill the baby”) and a cognitive behavioral representation (or a response that considers saving more lives at the expense of sacrificing a baby). What is interesting about this interpretation is that he stresses the role of cognition and ignores the role of currency-like emotions. Unlike other dilemmas, Greene thinks that, in this dilemma, the cognitive response “effectively competes” with the strong emotional response because a thorough “cost-benefit analysis” reveals that smothering the baby is the favorable option. “After all,” Greene reasons, “the baby is going to die no matter what...so you have nothing to lose...and much to gain by smothering [the baby], awful as it is” (p. 45). In other words, in this dilemma there is an unprecedentedly strong cognitive motivation. Before committing to Greene’s interpretation of the crying baby dilemma, I believe we should reexamine his interpretation and empirical results with the hope that Greene shows us the same consistency and evidence as in the footbridge dilemma.

Problems with Greene's Interpretation. Greene's interpretation is simplistic. It dramatizes the role of cognition and, at the same time, ignores the role of emotion when making consequentialist judgments. This is problematic since Greene is only singling out alarm bell emotions. Even if one tried to circumvent the problem by focusing solely on his theoretical commitments, the lack of evidence weighs in favor of a cooperative and interdependent, not a combative and independent interpretation of both processes. For, relative to other dilemmas, in this dilemma subjects experience a distinctive combination of high emotional and cognitive activity (Greene, 2008). Though subjects who endorse the consequentialist judgment do, overall, show more cognitive activity than subjects who endorse the deontological judgment (Greene, Nystrom, Engell, Darley & Cohen, 2004), Greene must be able to associate cognition only with consequentialist judgments. To do so, Greene would have to supplement his neuro-images with RT evidence (just as he did in his personal dilemma). Unfortunately, there is no RT evidence. Subjects, Greene tells us, take an equally long time to endorse either option (Greene, Sommerville, Nystrom, Darley & Cohen, 2001; Cushman, Young, & Greene, 2010). Given the lack of evidence, we have no reason to believe that two systems are in competition. I will try to offer a better interpretation.

A Theoretically Consistent Interpretation. Greene thinks that in theory consequentialist judgments are motivated by currency-like emotions. Yet, when interpreting the crying baby dilemma, he doesn't mention their role. I will piece together a much more consistent picture—one that takes into consideration the role of emotion.

One could surely come up with an interesting list of possibilities. Let's imagine a person, say, Maria Guadalupe who is motivated by four different kinds of relationships. She could be motivated to save (1) herself (i.e., a selfish emotion), (2) her community (i.e., prosocial emotion), (3) her family (i.e., familial emotion) or (4) her offspring (attachment based emotion).

I will begin with the first. Maria is hiding from enemy soldiers in a basement with her fellow townspeople. The soldiers are approaching the basement and her baby begins to cry. Automatically she puts her hands on the baby's mouth, being forced to decide whether to smother her baby (thus, saving the village members) or momentarily spare the baby's life (thus, placing everyone's life in jeopardy). Though compelled to save the baby at first, Maria realizes that her own life is also at risk. Maria is saddened because she is so young (mid-twenties) and hasn't accomplished anything on her bucket list. Weighing the pros and cons, she ultimately concludes that she would rather save her own life rather than the baby's. However selfish, her "cost-benefit analysis" convinces her that she can always have another baby.

Altruism or prosocial behavior is, perhaps, what Greene has in mind when offering his interpretation. Imagine the conditions are the same as in the last scenario except that, before deciding, Maria takes into consideration the townspeople, which may include anyone from her mailman to her childhood best friend. Since Maria has been an active member of the community, she has formed close relationships with most people in the basement. Ultimately, she sympathizes with her community and reasons that it would

be selfish of her to risk the lives of everyone she knows and loves for the sake of her beloved baby.

The third option is like the second, though her care and concern does not extend beyond her family. Imagine the conditions are the same as in the last scenario except that Maria considers her family members before deciding what action to endorse. Since Maria sympathizes with her grandparents, uncles, aunts, cousins, mother, father, and siblings, eventually a cost-benefit analysis convinces her to save all her family members instead of her baby.

The final option is like the third, except that her care and concern is confined to her children. Imagine the conditions are the same as in the last scenario except that Maria has two kids. In addition to her baby, she has a four-year old child who is among the townspeople. She sympathizes with her children and reasons that more pain will befall her older child because he is more conscientious of his surroundings. Plus, she tells herself, saving her baby may incur more future pain because the soldiers may be cruel executioners. Eventually, her thorough cost-benefit analysis persuades her to save her older child.

In all these cases emotions are involved in different ways, but yield the same consequentialist judgment. Though Greene may only agree with the second scenario (prosocial emotion or community based response), the rest of the scenarios are just as consistent with the crying baby dilemma because one can interpret “townspeople” or “village” members in multiple ways. That is, since Greene allows subjects to imagine

whom to include among their list of townspeople or village members, each subject may uniquely envision the victims whose lives will be at stake. Depending on the relationships involved, the emotional responses will also differ.

Given the multiplicity of different emotional possibilities, we must reinterpret this dilemma. In passing one may have interpreted the crying baby dilemma as a situation forcing subjects to choose between remote village members and a beloved infant. The former option being the impartial, neutral, cognitive approach; the latter being the partial, biased, emotionally involved approach. Attempting to create a much more consistent picture shows us that subjects are not necessarily torn between choosing a cognitively appealing option and an emotionally appealing option. Both options are just as emotionally and motivationally appealing. Elsewhere, Greene et al. (2004) tell us that the cognitive interpretation that was initially provided is not entirely accurate:

The account we've offered is complicated by the fact that the brain regions other than the DLPFC and inferior parietal lobes predict utilitarian judgment. One of these regions is in the posterior cingulate, which has been associated with emotion...This finding does not necessarily undermine [our view]. It does, however, challenge the overly simplistic view that utilitarian judgments are wholly allied with "cognition" while nonutilitarian judgments are wholly allied with "emotion." (p. 397)

In other words, the idea that consequentialist judgments (i.e., utilitarian judgments) are "wholly allied" with cognition overlooks the fact that emotions are also allied with the

same judgment. Given his theoretical commitments, I have tried to give an account of possible emotions and relationships that may motivate a subject to endorse the consequentialist judgment. In speculating possible emotions, however, another problem emerges, which I call Greene's fork. Is there really a difference between currency-like emotions and alarm bell emotions in these mentioned examples? Is it true that alarm-bell emotions are incompatible with cognition?

Greene's Fork. According to Greene, no alarm bell emotion can ever motivate a subject to endorse the consequentialist judgment. Similarly, no currency-like emotion can ever motivate subjects to endorse the deontological judgment. I think this fork is misleading. It is quite possible for the very same emotion to motivate the consequentialist and deontological judgment, even in the same moral situation. The first example will stress the role of emotions. The second example will stress the role of cognition. I will argue for the former by grounding a subject's willingness to endorse the saving-village option in an attachment-based emotion. Before arguing for my first point, I should mention that Greene takes the baby-saving option to be motivated by an alarm bell emotion. Discussions that explain the "building blocks" of moral emotions may support Greene's belief that contemplating infanticide would spawn a primitive, emotional response. Jorge Moll et al. (2008) incorporate attachment as one of the three building blocks of moral emotions (aggression and social rank being the others) because it "provides the basic ingredient for [inter-individual] bonding and affiliative behaviors, such as mother-offspring ties" (p. 7). Attachment, I believe, is a good candidate for

explaining why some acts, such as infanticide, evoke emotionally alarming responses. To require people to violate a deeply personal connection, i.e., like parent-offspring relationship, goes against one of the most fundamental relationship-sustaining emotions. Hence, it seems very likely that part of the reason why baby-saving option triggers a powerful alarm bell emotion is because it is grounded in this primitive, attachment-based emotion. I will try to show that the consequentialist judgment can also be driven by the same attachment-based emotion.

Let's reexamine the example of Maria Guadalupe and her children. Like the original scenario, her baby begins to cry as the soldiers are approaching, thus, being forced to decide between smothering her baby or risk being caught. In this case, note that one of Maria's children is also among the townspeople. Consequently, she is pinned against two identical attachment-based emotions. More specifically, she is torn between two competing behavioral representations, commanding her to rescue them both (don't let them die!). But, according to the dilemma, she is unable to consistently perform both options. Presumably, Maria would, at first, be paralyzed between these two options because there is no easy "cost-benefit analysis" or neutral or unbiased approach to help her decide between saving her infant or saving her other child (plus everyone else in the village). Ultimately, whatever option Maria endorses, the motivation will be the same.

The second example is like the first, but the only difference is that Maria has only one child, a baby. Let's suppose that Greene is right to say one option is more cognitively engaging, for surveying the entire situation reveals that the baby will die not matter what

option is endorsed. So, the most consequentialist thing for Maria to do, in this case, is to save her fellow townspeople. Yet, it is possible for Maria to reluctantly smother her baby while, at the same time, only have her baby's interest in mind. Perhaps, Maria hates her townspeople. But she, nevertheless, indirectly saves them because she realizes that leaving her baby at the hands of merciless, enemy soldiers might turn out to be a much more tragic death for her baby. Thus, she performs the act herself while all along hating her townspeople. Again, in this example, an attachment-based emotion coupled with some good o' cost-benefit analysis drove Maria to endorse the consequentialist judgment (though, quite honestly, she doesn't care about the numbers).

Given these two examples, we should not be reluctant to conclude that alarm bell emotions can drive consequentialist judgments. In the first example, I focused solely on alarm bell emotions to show that the same emotion can motivate both judgments. Thus, Greene's claim that each judgment has its own "fundamentally different" emotion is questionable. In the second example, I focused on cognition to show that it is compatible with an alarm bell emotion. Thus, Greene's belief that cognition and alarm bell emotions are incompatible is also questionable. If I am correct in my interpretation, these two examples show that the sharp fork is misleading and, perhaps, illusory. More importantly, it shows that consequentialist judgments can be driven by alarm bell emotions.

Objections

One could argue that I am ignoring the evolutionary rationale which underlies Greene's hypothesis. As the consequence, one could add, I am omitting the features that

comprise his personal factor. Including these features may be able to explain why some subjects endorse the atypical judgments. I disagree. Even if we consider Greene's evolutionary story, which explains why certain actions are emotionally salient, I still think we will run into the same difficulties. The problem, I add, is not the absence of the evolutionary story, but the absence of clarity in Greene's personal-impersonal distinctions. Even Greene acknowledges that he may need to modify his personal-impersonal distinction in a footnote (Greene, 2008, p.77n2).

One could also argue that I am holding the evidential standard much too high. Since every hypothesis has anomalies, we can expect that there will always be nonconforming subjects who endorse the atypical response. Though I am sympathetic to Greene's hypothesis (without the need for RT evidence), I find that his intent to reverse historical stereotypes and aim to establish his strong metaphysical claim requires strong evidence. Greene, in my opinion, needs to convince us that systems are, indeed, competitive and do yield different solutions. Further, he needs to convince us that our moral judgments map onto these different approaches or ways of thinking. Since the implications are groundbreaking, we need to rely on more than just the bare minimum.

One could also disagree with my arguments because they exaggerate the role of psychological processes. Like Greene, I am exaggerating the role of one psychological process as the expense of the other. For instance, in the switch dilemma, I emphasized the role of cognition and deemphasized the role of emotion when subjects made the deontological judgment. Also, in the crying baby dilemma, I emphasized the role of

emotion and deemphasized the role of cognition when subjects were making the consequentialist judgment. I think Greene and I can meet somewhere in the middle. All judgements may require a unique emotional-cognitive ratio (assuming these distinctions are real). Admitting this last point, however, undercuts [M] because the nature of each judgment no longer depends on combative, but cooperative systems. Thus, I add, there is a hidden psychological complexity.

One could object that my second argument unfairly and unrealistically portrays emotions. Perhaps, the only realistic and charitable emotion is the prosocial emotion because it captures what Greene means by consequentialist thinking, i.e., saving more lives at the expense of one. In other words, someone may complain that it is improbable that subjects would imagine such unrealistic examples. Like Greene, I think that each dilemma invites subjects to imagine themselves in the scenarios (Cushman & Greene, 2012). How one imagines the scenario, however, depends on the dilemma's configuration (i.e., the language being used). In the crying baby dilemma, Greene et al. (2004) use relational language (i.e., your townspeople and your baby), which differs from other dilemmas' relationally distant language (i.e., workers and large stranger). The differences in noun usage, I add, allows for interpretive variability. For, if we think about it, John Doe could sensibly consider anyone from his close-knit village, ranging from his local dentist to high school math teacher, from his work acquaintance(s) to childhood best friend, from his pet hamster to pet Chihuahua, from his siblings to parents, and from his spouse to children. Given that one could inclusively sympathize with all members or

exclusively sympathize with some members or none, all the emotions and scenarios mentioned are viable possibilities. Hence, I don't think I am unfairly or unrealistically characterizing emotions.

Yet, one could add that the principle of charity requires that I only consider the most consistent interpretation. Thus, I should only consider the prosocial emotion in my second argument. Thus, if considered a prosocial emotion, then perhaps it is not an alarm bell emotion. Even if we allow for this interpretation, I would argue that prosocial emotions should still be considered as an alarm bell emotion. First, one could easily come up with an evolutionary rationale that explains why prosocial emotions are antiquated responses (i.e., they are triggered by the prospect of harm towards a group one closely identifies with). Thus, if it is antiquated response, then it is also an alarm bell emotion. Second, prosocial emotions may only be effective because of the cross-relational attachments formed within members of a community. Even though in general, prosocial emotions may benefit the majority, there is a sense in which its influence or motivational force is grounded in the capacity to form attachments. Since there are some attachment-like elements grounding prosocial emotions, it seems reasonable to classify these as alarm bell emotions.

Implications/Conclusion

My thesis directly challenges Greene's simplistic account of moral judgments and suggest there is a hidden psychological complexity among all made judgments. His stronger thesis holds that [1] deontological judgments are driven by alarm bell emotions, and that [2] consequentialist judgments are driven by cognition (and currency-like emotions). I argued that Greene's evidence and theoretical commitments are consistent with the opposite of claims, namely, that [3] deontological judgments are driven by cognition, and [4] consequentialist judgments are driven by alarm bell emotions. Given that the opposite set of claims can also be held, Greene may be arguing beyond what the evidence directly allows. Given that his evidence and theoretical commitments are intended to establish [M], my counter set of claims undercuts this strong metaphysical claim. Perhaps moral judgments are not manifestations of psychological natural kinds. Perhaps it makes more sense to say that judgments are composed of both psychological elements.

Also, my arguments take issue with the simple characterizations of moral judgments and their relationship to discrete psychological processes. For instance, Greene thinks that engaging in a "cost-benefit analysis" directly relates to the normative concern (saving more lives at the expense of one). As I showed in some of the Maria Guadalupe examples, this normative concern is not always what motivates subjects to endorse the consequentialist option. In some cases, people can act selfishly or act in ways that only benefit those they directly care about, yet unintendedly help those who also happen to be

involved. It is wrong to think that consequentialist thinking or engaging in a “cost-benefit analysis” is by default the neutral, impartial, and altruistic approach. Thus, there is an incongruence between the normative concern and psychological properties.

Further my argument against Greene’s fork raises doubts about the way he draws his distinctions between emotions. As I showed, an alarm bell emotion can motivate two different moral judgments in the same moral situation. Also, I showed an attachment based emotion (presumed to be an alarm bell emotion) can function like a currency-like emotion. Though it functions as an imperative, for instance, urging Maria to save her child, a combined cost-benefit analysis may also urge Maria to save her child from a much more tragic death (i.e., it would be better for her child to die from loving hands rather than ruthless soldiers). Alarm bell emotions, therefore, are compatible with cognition and, moreover, can drive consequentialist judgments.

Lastly, Greene maintains that deontology is a kind of moral confabulation because its judgments are inspired by alarm bell emotions. This idea was primarily motivated by his strong metaphysical claim. Deontology, Greene tells us, is the outcome of very creative rationalizers, who are totally unaware of the intuitive and emotional processes fueling their judgments. The same, however, can be said for consequentialism. Subjects who make consequentialist judgments can also be prone to rationalizations because the same alarm bell emotions motivate them. Take, for instance, the case where Maria is pinned between two attachment-based emotions. Though Maria may unconsciously decide to save her older child at the expense of losing her baby, she may later tell herself,

friends, and family that she did it for the sake of everyone else. She may even tout consequentialist mantras while all along unaware of her real motive. If we consider Greene as an advocate for consequentialist decision-making (like Baron), then my emotional interpretation of consequentialist judgments show that this set of judgments are also inspired by the same emotion that allegedly fueled Kant's moral philosophy. Ironically, then, one could say that Greene's soul is not that different from Kant's after all. Both moral judgments, as Haidt would radically claim, can be driven by emotional intuitions and can be liable to rationalizations.

References

- Baron, J. (1994). Nonconsequentialist decisions. *Behavioral and Brain Sciences*, 17(01), 1-10.
- Cushman, F. (2013). Action, outcome, and value a dual-system framework for morality. *Personality and social psychology review*, 17(3), 273-292.
- Cushman, F., Young, L., & Greene, J. D. (2010). Our multi-system moral psychology: Towards a consensus view. *The Oxford handbook of moral psychology*, 47-71.
- Cushman, F., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social neuroscience*, 7(3), 269-279.
- Foot, P. (1967). The problem of abortion and the doctrine of double effect.
- Greene, J. D. (2007). Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains.
- Greene, J. D. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 3). Cambridge: MIT Press.
- Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1451), 1775.

- 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*, *293*(5537), 2105-2108.
- Greene, J. D., & Haidt, J. (2002). How (and where) does moral judgment work?. *Trends in cognitive sciences*, *6*(12), 517-523.
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, *44*(2), 389-400.
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*(3), 1144-1154.
- Haidt, J. (2001). The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological review*, *108*(4), 814.
- Haidt, J., Bjorklund, F., & Murphy, S. (2000). Moral dumbfounding: When intuition finds no reason. *Unpublished manuscript, University of Virginia*.
- Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological bulletin*, *109*(2), 163.

- Moll J., de Oliveira-Souza R., Zahn R., Grafman J. (2008) The cognitive neuroscience of moral emotions. In W. Sinnott-Armstrong (Ed.), *Moral Psychology* (Vol. 3). Cambridge: MIT Press.
- Paxton, J. M., Ungar, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, 36(1), 163-177.
- Piaget, J. (1965). *The moral judgement of the child*. (M. Gabain, Trans.) New York: Free Press. (Original work published 1932).
- Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 12, 643–662.
- Thomson, J. J., & Parent, W. (1986). *Rights, restitution, and risk: Essays in moral theory*. Harvard University Press.
- Turiel, E. (1983). *The development of social knowledge: Morality and convention*. Cambridge University Press.
- Tiberius, V. (2014). *Moral psychology: A contemporary introduction*. Routledge.