

Experimental Challenges to Intuition, Pt. 2: fMRI Studies

I. Greene et al.'s Empirical Findings

Two classes of particular-case moral judgments:

1. *characteristically consequentialist judgments*: Particular-case moral judgments that are “easily justified in terms of the most basic consequentialist principles” (“Secret Joke,” p. 6).

Ex.: One’s judgment that it is permissible to turn the switch in the *bystander trolley case*.

2. *characteristically deontological judgments*: Particular-case moral judgments that are “in favor of characteristically deontological conclusions” (*ibid.*, p. 5).

Ex.: One’s judgment that it is impermissible to push the fat man in the *footbridge trolley case*.

Two kinds of psychological processes:

1. “*cognitive*” processes: Information processing that involves “inherently neutral representations . . . that do not automatically trigger particular behavioral responses or dispositions” (*ibid.*, p. 7).

“Cognitive” representations are *highly flexible*: they “can be easily mixed around and recombined as situational demands vary, and without pulling the agent in sixteen different behavioral directions at once” (*ibid.*, p. 7).

“Cognitive” processes are “especially important for reasoning, planning, manipulating information in working memory, controlling impulses, and ‘higher executive functions’ more generally” (*ibid.*, p. 8).

Regions of the brain associated with “cognitive” processes:

- *the dorsolateral prefrontal cortex*;
- *the dorsolateral parietal lobe*.

2. *emotional processes*: Information processing that involves behaviorally valenced representations that trigger automatic effects and hence “have direct motivational force” (“Neural Bases,” p. 397).

Emotional processes are *very speedy*: “emotions are very reliable, quick, and efficient responses to re-occurring situations” (“Secret Joke,” p. 33).

Regions of the brain associated with emotional processes:

- *the amygdala*;
- *the medial frontal cortex*;
- *the medial parietal lobe*.

Greene et al.’s hypothesis: Characteristically deontological judgments tend to be driven by emotional processes, while characteristically consequentialist judgments tend to be driven by “cognitive” processes.

This hypothesis yields several predictions:

prediction #1: The contemplation of cases like the footbridge trolley case should produce increased neural activity in brain regions associated with emotional processes.

prediction #2: The contemplation of cases like the bystander trolley case should produce increased neural activity in brain regions associated with “cognitive” processes.

prediction #3: People who reach a non-standard verdict about cases like the footbridge trolley case should take longer to reach a verdict (since they are overriding an emotional response).

prediction #4: People who reach a non-standard verdict about cases like the bystander trolley case should take approximately as long to reach a verdict.

In order to test these predictions, Greene et al. appealed to a “purely descriptive” way of sorting moral dilemmas that are “like the footbridge trolley case” (and hence give rise to deontological judgments) from those that are “like the bystander trolley case” (and hence give rise to consequentialist judgments): the former types of cases were deemed to involve *harm that is brought about in an “up close and personal” way*, whereas the latter were deemed to involve *harm that is brought about in an impersonal way*.

This yields two sorts of moral dilemmas:

1. *personal moral dilemmas:* Moral dilemmas, like the footbridge trolley case, in which one option involves harming someone in an “up close and personal” way.

Greene et al. propose the following “first cut” attempt at characterizing what makes a moral dilemma personal: “the action in question (a) could reasonably be expected to lead to serious bodily harm (b) to a particular person or a member or members of a particular group of people (c) where this harm is not the result of deflecting an existing threat onto a different party” (“An fMRI Investigation,” p. 2107, n. 9).

2. *impersonal moral dilemmas:* Moral dilemmas, like the bystander trolley case, that do not fall in the above category.

Using this as their criterion for deeming a case to be “like the footbridge trolley case” or “like the bystander trolley case,” Greene et al. used *functional Magnetic Resonance Imaging* (fMRI) to see which parts of the brain were active (in the sense of having blood flow through them) when various subjects considering moral dilemmas in these two categories. Their findings: predictions #1–4 were (for the most part) confirmed.

(Why “for the most part”? The mean reaction time for people reaching non-standard verdicts about impersonal moral dilemmas was slightly higher than that for people reaching standard verdicts about those cases. Also, activity in the *posterior cingulate*, a region of the brain that has been associated with emotion, was detected during contemplation of impersonal moral dilemmas [i.e. dilemmas deemed to give rise to consequentialist judgments]; cf. “Neural Bases,” p. 397.)

Two worries:

- Greene’s et al.’s tentative criteria for being a personal moral dilemma is inadequate.
- The personal vs. impersonal moral dilemma distinction—however one characterizes it—doesn’t seem to track the dilemma-giving-rise-to-a-deontological-judgment vs. dilemma-giving-rise-to-a-consequentialist-judgment distinction.

Most famously, a case in which there is a trapdoor under the fat man that you can trigger from afar gives rise to a characteristically deontological judgment, but is impersonal.

Also: Kamm’s lazy susan case, Nichols and Mallon’s teacups version of the fat man case.

For the sake of argument, let us (temporarily) grant to Greene that there is some way of characterizing the gives-rise-to-a-deontological-judgment/gives-rise-to-a-consequentialist-judgment distinction that confirms his hypothesis.

Moreover, Greene thinks there is a host of additional empirical evidence for his hypothesis, including Small and Lowenstein’s research on the “determinate victim effect,” neuroimaging studies of the desire to punish, Haidt’s research on moral condemnation of harmless actions, data on the moral judgments of lesion victims with emotional deficits, and studies of responses to trolley cases after one has been emotionally primed.

II. The Argument from Evolutionary History

Singer and Greene draw quite radical normative implications from Greene et al.'s empirical findings.

They begin by telling a “just so” story about the evolution of our faculty for making deontological judgments about personal moral dilemmas:

“For most of our evolutionary history, human beings have lived in small groups In these groups, violence could only be inflicted in an up-close and personal way To deal with such situations, we have developed immediate, emotionally based responses to questions involving close, personal interactions with others” (Singer, pp. 347–348).

In light of their appeal to such an evolutionary story, it is very tempting to read Singer and Greene as making something like the following argument (though I don't believe that they actually argue in this way):

the argument from evolutionary history:

1. Deontological intuitions are evolutionary by-products that were adapted to handle situations that we no longer find ourselves in. [*premise*]
2. So, deontological intuitions, unlike consequentialist intuitions, do not have any genuine normative force. [*follows from 1*]

This is a bad argument. Presumably consequentialist intuitions are just as much an evolutionary by-product as deontological intuitions, so this gives us no reason to privilege consequentialist intuitions.

Singer contends that consequentialist intuitions “do not seem to be . . . the outcome of our evolutionary past,” but I find this claim completely astounding (p. 350).

Moreover, Singer and Greene are committed to mathematical and scientific judgments having genuine normative force, yet presumably our faculty for making these judgments has an evolutionary basis.

Singer calls for us to engage in “the ambitious task of separating those . . . judgments that we owe to our evolutionary and cultural history, from those that have a rational basis” (p. 351), yet clearly this is a false dichotomy.

III. The Argument from Morally Irrelevant Factors

Closer to an argument that Singer and Greene actually make is the following:

the argument from morally irrelevant factors:

1. The emotional processing that gives rise to deontological intuitions responds to factors that make a dilemma personal rather than impersonal. [*premise*]
2. The factors that make a dilemma personal rather than impersonal are morally irrelevant. [*premise*]
3. So, the emotional processing that gives rise to deontological intuitions responds to factors that are morally irrelevant. [*follows from 1, 2*]
4. So, deontological intuitions do not have any genuine normative force. [*follows from 3*]

Some worries about this argument:

- In order to establish premise 1, Singer and Greene need it to be the case that the personal vs. impersonal dilemma distinction really does track the deontological vs. consequentialist judgment distinction. However, this is problematic, given that (i) Greene's current way of drawing the personal vs. impersonal distinction is not up to the task, and (ii) formulating a principle that distinguishes what separates cases-eliciting-a-deontological-judgment from cases-eliciting-a-consequentialist-judgment is likely to be as difficult as the original problem of formulating a principle that distinguishes the permissible trolley cases from the impermissible trolley cases.

- Premise 2 is the crucial step in the argument. However, it appears to be defended *from the armchair* by means of an *appeal to intuition*. Moreover, once we eventually have our hands on the right way of drawing the cases-eliciting-a-deontological-judgment vs. cases-eliciting-a-consequentialist-judgment distinction, it is far from clear that every feature picked up by the faculty eliciting deontological reactions but ignored by the faculty eliciting consequentialist reactions will, intuitively, be morally irrelevant. (For example, on Greene's own way of drawing the line, the faculty eliciting consequentialist reactions seems to be ignoring one morally salient feature, namely the separateness of persons—the fact that people are not fungible commodities that can be interchanged willy-nilly.)
- Indeed, the epistemic efficacy of consequentialist vs. deontological intuitions now seems to be purely a function of *what sorts of features out there in the world they are responding to*. Thus issues about the evolutionary history of our dispositions to have these intuitions, and issues about the “alarm-like” and “emotion-based” nature of deontological intuitions, and even issues about the fact that these two sets of intuitions stem from separate faculties, seem to have dropped out of the picture.

IV. The Argument from Moral Rationalization

Let us suppose that the previous argument actually works. Then people in the applied ethics literature who appeal to distinctively non-consequentialist intuitions about cases are in trouble, but many classical deontologists are untouched by the argument as presented so far.

For example, Kant's arguments for the Categorical Imperative do not appeal to moral intuitions about cases (of either a deontological or a consequentialist variety). Indeed, you might think that this feature is one of the chief attractions of Kant's moral philosophy.

However, Greene extends his argument so that even deontologists who do not explicitly appeal to intuitions fall under his attack.

Let us assume (somewhat implausibly) that deontologists end up defending theories that endorse almost all of their emotionally driven moral reactions, even if they don't in fact appeal to those reactions when defending their theories. Then Greene appears to argue as follows:

the argument from moral rationalization:

1. Deontologists defend theories that endorse almost all of their emotionally driven deontological intuitions about cases. [*premise*]
2. One explanation for (1) is that deontologists are drawn to their moral theories by their emotional reactions, and then construct a *post hoc* rationalization for their theories as a sort of “moral confabulation.” [*premise*]
3. Another explanation for (1) is that our emotionally driven deontological intuitions about cases happen to track the rationally discoverable moral truth. [*premise*]
4. Given the argument from morally irrelevant factors, it would be an amazing coincidence if our emotionally driven deontological intuitions about cases happen to track the rationally discoverable moral truth. [*premise*]
5. So, the explanation in (2) is a better explanation of (1) than the explanation in (3). [*follows from 2, 3, 4*]
6. So, we have good reason to believe that deontological philosophy is an exercise in moral rationalization. [*follows from 1, 5 by inference to the best explanation*]

Clearly this argument depends very heavily on the previous argument.

Another worry: even if we did have good reason to think that deontological moral theories are wrong, we would still want to know *where* they go wrong. The goal in philosophy is not to show people up by proving *that* they're wrong, but rather to understand *why* they're wrong (if indeed they are).