Georgetown University Law Center

Scholarship @ GEORGETOWN LAW

2011

Emotion, Neuroscience, and Law: A Comment on Darwin and Greene

John Mikhail
Georgetown University Law Center, jm455@law.georgetown.edu

This paper can be downloaded free of charge from:
https://scholarship.law.georgetown.edu/facpub/611
http://ssrn.com/abstract=1761295

Emotion, Neuroscience, and Law: A Comment on Darwin and Greene

John Mikhail

Abstract: Darwin's (1871) observation that evolution has produced in us certain emotions responding to right and wrong conduct that lack any obvious basis in individual utility is a useful springboard from which to clarify the role of emotion in moral judgment. The problem is whether a certain class of moral judgments is "constituted" or "driven by" emotion (Greene 2008, p. 108) or merely correlated with emotion while being generated by unconscious computations (e.g., Huebner et al. 2008). With one exception, all of the "personal" vignettes devised by Greene and colleagues (2001, 2004) and subsequently used by other researchers (e.g., Koenigs et al. 2007) in their fMRI and behavioral studies of emotional engagement in moral judgment involve violent crimes or torts. These studies thus do much more than highlight the role of emotion in moral judgment; they also support the classical rationalist thesis that moral rules are engraved in the mind.

In *The Descent of Man*, Darwin (1981/1871, p. 70) affirmed his belief in an innate moral faculty, explaining that he fully agreed with Kant and other writers "that of all the differences between man and the lower animals, the moral sense or conscience is by far the most important." Darwin insisted that the moral sense is not a mysterious gift of unknown origin, however, but the natural result of evolution, with antecedents in the social instincts of other animals. He thus famously argued that "any animal whatever, endowed with well-marked social instincts, would inevitably acquire a moral sense or conscience, as soon as its intellectual powers had become as well developed, or as nearly developed, as in man" (id. at 71-72). And he laid the foundation of subsequent research on the evolution of morality by examining a range of animal traits and behaviors, including their sociability, desire for companionship, and the misery they feel when they are abandoned; their love, sympathy, and compassion for one another; and their mutual willingness to sacrifice themselves and to render services to one another when hunting or defending against attack.

Darwin held that the social instincts of nonhuman animals developed "for the general good of the community," which he defined as "the means by which the greatest possible number of individuals can be reared in full vigor and health, with all their faculties perfect, under the conditions to which they are exposed" (id. at 97-98). The same was true of *homo sapiens*, he inferred; therefore, neither egoism nor a universalistic hedonism (the "Greatest Happiness Principle") was descriptively adequate: "When a man risks his life to save that of a fellow-creature, it seems more appropriate to say that he acts for the general good or welfare, rather than for the general happiness of mankind" (id. at 98). Darwin endorsed Herbert Spencer's conclusion that "'the experiences of utility organized and consolidated through all past generations of the human race, have been

producing corresponding modifications, which, by continued transmission and accumulation, have become in us certain faculties of moral intuition—certain emotions responding to right and wrong conduct, which have no apparent basis in the individual experiences of utility'" (id. at 101-102). Finally, Darwin held that this combination of social instincts, intellectual powers, and effects of habit would "naturally lead to the golden rule: 'As ye would that men should do to you, do ye to them likewise.'" This rule, he averred, "lies at the foundation of morality" (id. at 106).

The idea that evolution has produced in us "certain emotions responding to right and wrong conduct" that lack any obvious basis in individual experiences of utility is a useful springboard from which to clarify an important problem in the cognitive science of moral judgment. The problem is how to understand the role of emotion in moral judgment, and specifically whether a certain class of moral judgments is "constituted" or "driven by" emotion (Greene 2008, p. 108; see also Greene 2004, 2009; Koenigs et al. 2007) or merely correlated with emotion while being generated by unconscious computations (e.g., Huebner et al. 2008). On at least some interpretations, there are important differences between these formulations, although these differences may disappear at certain neurocognitive or neurobiological levels of scientific description. My claim is that the second formulation—the Darwin-Spencer thesis, according to which emotions "respond to" independent moral appraisals—is a better working model of moral cognition with respect to this class of judgments.

To see why, it is useful to look closely at the 25 "personal" dilemmas devised by Greene and colleagues in their original fMRI study (2001) and subsequently used by a number of other researchers (e.g., Greene et al. 2004; Koenigs et al. 2007; Moore et al. 2008). Greene found that these vignettes elicited increased activity in the medial prefrontal cortex (MPFC), posterior cingulated cortex (PCC), superior temporal sulcus (STS), and amygdala. Because these regions are associated with emotional processing, he concluded that these "characteristically deontological" judgments are driven by emotion. What seems to have escaped his notice and that of the scientific community generally, however, is that *all of the actions described by these vignettes are well-known crimes or torts* (Table 1). Specifically, 22 of the 25 scenarios satisfy a prima facie case for purposeful battery and/or intentional homicide (i.e., murder). Two other cases involve acts of rape and sexual battery, while the final case describes a negligent (i.e. unreasonable) failure to rescue.

With one exception, then, what Greene actually did in the "personal" condition of his experiment was to put subjects in the scanner and ask them to respond to a series of violent crimes and torts. There are other relevant features of these scenarios, of course; some of them raise principal-agent problems and others involve duress or necessity, for example. Fundamentally, however, all of them describe acts that standard legal analysis would classify as serious wrongs, subject to conceivable, but ultimately weak, affirmative defenses. Moreover, all of them

involve serious bodily injury and thus implicate the right to physical safety. By contrast, only five of the 19 cases in Greene's "impersonal" condition are batteries, and only one of these batteries is purposeful. The other four cases involve foreseeable but non-purposeful harms, at least two of which admit of an uncontroversial necessity defense. The remaining 14 "impersonal" scenarios are a hodgepodge of cases that raise a variety of legal issues, including fraud, tax evasion, insider trading, public corruption, theft, unjust enrichment, and necessity as a defense to trespass to chattels. Finally, five of these residual cases describe risk-risk tradeoffs in the context of vaccinations and environmental policy.

The upshot is that Greene's (2001, 2004) original experiments did not really test two patterns of moral judgment—one "deontological" and the other "utilitarian"—as much as different categories of potentially wrongful behavior. The basic cleavage he identified in the brain was not Kant versus Mill, but purposeful battery, rape, and murder, on the one hand, and a disorderly grab bag of theft crimes, regulatory crimes, torts against non-personal interests, and risk-risk tradeoffs, on the other. Moreover, his finding that the MPFC, PCC, STS, and amygdala are recruited for judgment tasks involving purposeful battery, rape, and murder does not undermine the traditional rationalist thesis that moral precepts are engraved in the mind (e.g., Grotius 1625; Kant 1788; Leibniz 1705). On the contrary, Greene's evidence largely supports that thesis. Crimes and torts have *elements*, and the relevant pattern of intuitions is best explained by assuming that humans possess implicit knowledge of moral and legal *rules*. Naturally, violent crimes and torts are more emotionally engaging than insider trading or environmental risk analysis, but it does not follow that emotion "constitutes" or "drives" the judgment that the former acts are wrong. Rather, what drive these intuitions are the unconscious computations that characterize these acts as battery, rape, or murder in the first place. By mischaracterizing their own stimuli, then, Greene and other neuroscientists (e.g., Koenigs et al. 2007) have drawn specious conclusions and misconceived the nature of the problem.

Returning to Darwin, the main questions for cognitive science going forward include (1) how the brain computes unconscious representations of purposeful battery, rape, murder, negligence, and other forms of harmful trespass, and (2) how these computations and the negative emotions they typically elicit are related to the complex cognitive and socio-emotional capacities that humans share with other animals (cf. Darwin 1981/1871; Spencer 1978/1897; see generally Mikhail 2007, 2009, in press). Future research should focus more squarely on these topics and move beyond potentially misleading pseudo-problems such as how reason and emotion "duke it out" in the brain.

References

Darwin, C. (1981/1871). *The Descent of Man, and Selection in Relation to Sex*. Princeton, NJ: Princeton University Press.

Greene, J. (2008). The secret joke of Kant's soul; and Reply to Mikhail and Timmons. In W. Sinnott-Armstrong (Ed.), *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development*. Cambridge, MA: MIT Press, pp. 35-79, 105-117.

Greene, J., Sommerville, R., Nystrom, L., Darley, J., & Cohen, J. (2001). An fMRI investigation of emotional engagement in moral Judgment. *Science, 293*, 2105–2108.

Greene, J., Nystrom, L., Engell, A., Darley, J., & Cohen, J. (2004). The neural basis of cognitive conflict and control in moral judgment. *Neuron, 44*, 389-400.

Grotius, H. (1925/1625). *On the Law of War and Peace*. (F. W. Kelsey, Trans.). Oxford: Clarendon Press.

Huebner, B., Dwyer, S. & Hauser M. (2008). The role of emotion in moral psychology. *Trends in Cognitive Sciences, 13*, 1-6.

Kant, I. (1993/1788). *Critique of Practical Reason*. (L.W. Beck, Trans.). MacMillan: New York.

Koenigs, M., Young, L., Adolphs, R., Tranel, D., Cushman, F., Hauser, M., & Damasio, A. (2007). Damage to ventromedial prefrontal cortex increases utilitarian moral judgments. *Nature, 446*, 908–911.

Leibniz, G. (1981/1704). *New Essays on Human Understanding*. (P. Remnant & J. Bennett, Eds.). Cambridge: Cambridge University Press.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Sciences, 11*, 143-152.

Mikhail, J. (2009). Moral grammar and intuitive jurisprudence: A formal model of unconscious moral and legal knowledge. In B. H. Ross (Series Ed.) & D. M. Bartels, C. W. Bauman, L. J. Skitka, & D. L. Medin (Eds.), *Psychology of Learning and Motivation, Vol. 50: Moral Judgment and Decision Making*. San Diego, CA: Academic Press.

Mikhail, J. (in press). *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. Cambridge: Cambridge University Press.

Moore, A., Clark, B., & Kaine, M. (2008). Who shall not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science*, 19, 549-557.

Spencer, H. (1978/1897). *The Principles of Ethics*. Indianapolis: Liberty Classics.

Dilemma	Standard Legal Analysis
1. Transplant	Battery/Homicide
2. Footbridge	Battery/Homicide
3. Country Road	Negligent Failure to Rescue
4. Architect	Battery/Homicide
5. Lifeboat	Battery/Homicide
6. Hard Times	Rape/Sexual Assault and Battery
7. Smother for dollars	Battery/Homicide
8. Safari	Battery/Homicide
9. Crying Baby	Battery/Homicide
10. Plane Crash	Battery/Homicide
11. Hired Rapist	Rape/Sexual Assault and Battery
12. Grandson	Battery/Homicide
13. Infanticide	Battery/Homicide
14. Preventing the Spread	Battery/Homicide
15. Modified Lifeboat	Battery/Homicide
16. Modified Preventing the Spread	Homicide
17. Modified Safari	Battery/Homicide
18. Modified Bomb	Battery/Torture
19. Submarine	Battery/Homicide
20. Lawrence of Arabia	Battery/Homicide
21. Sophie's Choice	Battery/Homicide
22. Sacrifice	Homicide
23. Vitamins	Battery
24. Vaccine Test	Battery/Homicide
25. Euthanasia	Battery/Homicide

Table 1: Standard Legal Analysis of Greene et al.'s (2001) "Personal" Dilemmas