

Why are VMPFC patients more utilitarian? A dual-process theory of moral judgment explains

Joshua D. Greene

Department of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138, USA

Koenigs, Young and colleagues [1] recently tested patients with emotion-related damage in the ventromedial prefrontal cortex (VMPFC) using moral dilemmas used in previous neuroimaging studies [2,3]. These patients made unusually utilitarian judgments (endorsing harmful actions that promote the greater good). My collaborators and I have proposed a dual-process theory of moral judgment [2,3] that we claim predicts this result. In a Research Focus article published in this issue of *Trends in Cognitive Sciences*, Moll and de Oliveira-Souza [4] challenge this interpretation.

Our theory aims to explain some puzzling patterns in commonsense moral thought. For example, people usually approve of diverting a runaway trolley that mortally threatens five people onto a side-track, where it will kill only one person. And yet people usually disapprove of pushing someone in front of a runaway trolley, where this will kill the person pushed, but save five others [5]. Our theory, in a nutshell, is this: the thought of pushing someone in front of a trolley elicits a prepotent, negative emotional response (supported in part by the medial prefrontal cortex) that drives moral disapproval [2,3]. People also engage in utilitarian moral reasoning (aggregate cost–benefit analysis), which is likely subserved by the dorsolateral prefrontal cortex (DLPFC) [2,3]. When there is no prepotent emotional response, utilitarian reasoning prevails (as in the first case), but sometimes prepotent emotions and utilitarian reasoning conflict (as in the second case). This conflict is detected by the anterior cingulate cortex, which signals the need for cognitive control, to be implemented in this case by the anterior DLPFC [Brodmann's Areas (BA) 10/46]. Overriding prepotent emotional responses requires additional cognitive control and, thus, we find increased activity in the anterior DLPFC when people make difficult utilitarian moral judgments [3].

More recent studies support this theory: if negative emotions make people disapprove of pushing the man to his death, then inducing positive emotion might lead to more utilitarian approval, and this is indeed what happens [6]. Likewise, patients with frontotemporal dementia (known for their 'emotional blunting') should more readily approve of pushing the man in front of the trolley, and they do [7]. This finding directly foreshadows the hypo-emotional VMPFC patients' utilitarian responses to this and other cases [1]. Finally, we've found that cognitive load selectively interferes with utilitarian moral judgment,

bolstering our claim that utilitarian judgments are preferentially supported by controlled cognitive processes (J.D. Greene *et al.*, unpublished).

Moll and de Oliveira-Souza [4] challenge the dual-process interpretation of the VMPFC lesion data in two ways. First, they claim that this theory can only be supported by a double dissociation between utilitarian VMPFC patients and anti-utilitarian DLPFC patients. Although such results would provide ideal evidence for this theory, it does not follow that a single dissociation provides no evidence. Only the dual-process theory specifically predicts that VMPFC patients will be unusually utilitarian across a range of difficult cases, and the inclusion of brain-damaged control patients demonstrates the anatomical selectivity of this effect. Second, Moll and de Oliveira-Souza observe that regions in the anterior DLPFC associated with utilitarian judgment [3] are damaged in some VMPFC patients, suggesting that they ought to be less utilitarian. This is an interesting observation, but it need not undercut the dual-process interpretation. We've argued that these anterior DLPFC regions (BA 10/46) exhibit increased activity during utilitarian judgment because of their role in controlling countervailing emotional responses [3]. But if VMPFC patients lack these emotional responses in the first place, then there is no need for control. VMPFC patients should exhibit utilitarian judgment so long as they are capable of utilitarian reasoning, which is likely subserved by more posterior regions of the DLPFC, as far back as BA 46 [2,3] and BA 44 [3].

As an alternative to the dual-process theory, Moll and de Oliveira-Souza attribute the VMPFC patients' utilitarian judgments to reduced 'prosocial moral sentiments' that 'emerge from integration, instead of conflict, between emotional and cognitive mechanisms'. VMPFC patients confronted with difficult moral dilemmas might not experience conflict, but it appears that healthy people do [3], and only a dual-process theory can explain why. Moll and de Oliveira-Souza contrast pro-social emotions with 'self-centered' emotions supported by the DLPFC and orbitofrontal cortex. This contrast could explain why VMPFC patients are more utilitarian when utilitarian judgments happen to be self-serving, but it does not explain why patients with frontotemporal dementia [7], normal patients experiencing positive emotion [6], and the present VMPFC patients [1] make more utilitarian judgments even when no selfish motivation exists, as in the pushing case described above. Ironically, patients with emotional deficits may, in some contexts, be the most pro-social of all.

Corresponding author: Greene, J.D. (jgreene@wjh.harvard.edu).
Available online 10 July 2007.

References

- 1 Koenigs, M. *et al.* (2007) Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* 446, 908–911
- 2 Greene, J.D. *et al.* (2001) An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108
- 3 Greene, J.D. *et al.* (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400
- 4 Moll, J. and de Oliveira-Souza, R. (2007) Moral judgments, emotions, and the utilitarian brain. *Trends Cogn. Sci.* 11, 319–321
- 5 Thomson, J. (1985) The trolley problem. *Yale Law J.* 94, 1395–1415
- 6 Valdesolo, P. and DeSteno, D. (2006) Manipulations of emotional context shape moral judgment. *Psychol. Sci.* 17, 476–477
- 7 Mendez, M.F. *et al.* (2005) An investigation of moral judgment in frontotemporal dementia. *Cogn. Behav. Neurol.* 18, 193–197

1364-6613/\$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tics.2007.06.004

Letters Response

Response to Greene: Moral sentiments and reason: friends or foes?

Jorge Moll and Ricardo de Oliveira-Souza

Cognitive and Behavioral Neuroscience Unit, LABS-D'Or Hospital Network, Rio de Janeiro, RJ 22281-080, Brazil

We appreciate Greene's comments on our article [1]. We have argued that neither simple emotional blunting (i.e. somatic markers) nor Greene's dual-process view of opposing emotional and cognitive mechanisms would satisfactorily explain the Koenigs *et al.* findings on patients with ventromedial prefrontal cortex (VMPFC) damage [2,3]. Although these patients made more utilitarian choices in trolley-type dilemmas (i.e. less emotional and more rational decisions), they opted more often for costly punishing non-cooperators in the ultimatum game (i.e. they were more emotional). Based on experimental and theoretical grounds [4,5], we contend that dissociation within the moral sentiment domain offers a more parsimonious explanation for these findings. In our view, the VMPFC and the frontopolar cortex (FPC), a region greatly expanded in humans [6], in conjunction with temporal and limbic-basal forebrain systems [5], play a distinguished role in the experience of prosocial sentiments (i.e. guilt, compassion and interpersonal attachment), whereas the ventrolateral PFC is more relevant for the experience of anger or indignation. Reduced prosocial sentiments are compatible with increased 'cold-blooded' utilitarian choices in personal dilemmas [3], and preserved or increased punishment of others in the ultimatum game [2].

Specifically regarding the moral judgment study by Koenigs *et al.* [3], we suggested that a single process — a decrease of certain emotions — would be sufficient to explain the effect of the increased utilitarian choices of VMPFC patients. In response, Greene argues that this would not disprove his dual-process theory. We agree with him on this point, but disagree that 'Only the dual-process theory specifically predicts that VMPFC patients will be unusually utilitarian across a range of difficult cases...' [1]. Although the dual-process theory is not incompatible with this effect, it does not provide a better explanation than a single process hypothesis of reduced emotions. The same argument applies to

evidence, cited by Greene, that utilitarian responses were increased in frontotemporal dementia and that they can be facilitated by mood induction [1]. Greene also points to unpublished data showing that cognitive load can modulate judgments; this is an interesting point, but a number of studies have shown that directed attention mediated by the PFC can reduce or enhance emotion experience (e.g. Ref. [7]), which again challenges the notion that cognitive control and emotion are essentially competitive mechanisms.

From an anatomical perspective, we pointed out that the brain damage of VMPFC patients in the study by Koenigs *et al.* [3] extended to the same FPC and anterior DLPFC regions activated by personal and utilitarian judgments in Greene's fMRI studies [8,9]. Therefore, one would have expected impairments of utilitarian reasoning, in addition to emotional impairments. In response, Greene suggests that the capacity for utilitarian reasoning could be sought for in more posterior DLPFC regions [Brodmann's Areas (BA) 44/46] [1]. However, these areas were not found to specifically support utilitarian choices in highly conflicting moral judgments [8]. Instead, these regions were non-specifically activated by low conflict moral (impersonal) and non-moral judgments [9]. These effects might reflect general executive, working memory, language and other functions mediated by posterior DLPFC and, therefore, can hardly be considered to be specific for 'utilitarian reasoning' or 'cognitive control' mechanisms.

To summarize, the seemingly paradoxical finding of increased rational or less emotional utilitarian judgments [3] and increased emotional or less rational choices in the ultimatum game [2] of VMPFC patients can neither be explained by overall emotional blunting nor by Greene's dual-process theory. By contrast, a selective impairment of prosocial sentiments, with preserved capacity for anger or indignation, can account for these findings. That VMPFC patients make more prosocial choices (from a utilitarian perspective) is a reminder of the gulf that divides observable behaviors and internal motivations. The apparently 'prosocial'

Corresponding author: Moll, J. (mollj@neuroscience-rio.org).
Available online 16 July 2007.