

DO ABNORMAL RESPONSES SHOW UTILITARIAN BIAS?

Guy Kahane and Nicholas Shackel

Neuroscience has recently turned to the study of utilitarian and non-utilitarian moral judgment. Koenigs *et al.*¹ examined the responses of normal and ventromedial-prefrontal-cortex damaged subjects to moral scenarios drawn from fMRI studies by Greene *et al.*^{2,3,4} and claim that patients with VMPC damage have an abnormally 'utilitarian' pattern of moral judgment. It is crucial to Koenigs *et al.*'s claims that Greene *et al.*'s scenarios pose a conflict between utilitarian consequence and duty; however many of them do not meet this condition. Because of this methodological problem it is too early to claim that VMPC patients have a utilitarian bias.

Greene *et al.* reported that brain areas typically associated with affect are activated when subjects make moral judgements about 'personal' scenarios where one alternative requires directly causing serious harm to persons. They found that in the minority that judges that such choices are appropriate, areas associated with cognition and cognitive conflict were activated as well. On the basis of a later study that found similar results in responses to 'difficult' personal scenarios, Greene has further suggested that the controversies between utilitarian and non-utilitarian views of morality 'might reflect an underlying tension between competing subsystems in the brain'⁴, a claim taken up by leading ethicists⁵.

Koenigs *et al.* drew on Greene *et al.*'s battery of moral scenarios to compare normal subjects with six subjects who have focal bilateral damage to the VMPC, a brain region associated with the normal generation of emotions and, in particular, social emotions. They report that these patients "produce an abnormally 'utilitarian' pattern of judgements on [personal] moral dilemmas... In contrast, the VMPC patients' judgements were normal in other classes of moral dilemmas."¹ These claims are based on VMPC patients' pattern of response to 'high-conflict' scenarios, a subset of personal scenarios on which normal subjects tended to disagree and which elicited greater response times.

However, the methodology used by Koenigs *et al.* cannot support claims about a utilitarian bias. Data from the categorization of the scenarios by five professional moral philosophers shows that many are not of the required type. Only 45% of their impersonal scenarios and 48% of the personal ones were classified as involving a choice between utilitarian and non-utilitarian options. Koenigs *et al.*'s distinction between low and high-conflict scenarios does not correspond to a difference in the scenarios' content. The high conflict scenarios are not all clear cases of utilitarian choice and some low conflict ones are very clear cases of such choice: Of the 13 high-conflict scenarios, our judges classified

only 8 as pure cases of utilitarian vs. non-utilitarian choice; conversely, two low-conflict scenarios *were* classified as such.

The battery of personal scenarios is therefore not an adequate measure of utilitarian choice and the distinction between low and high-conflict reflects only a difference in behavioural response rather than consistent differences in the content of the scenarios. Thus it is too early to claim that VMPC patients have a bias towards utilitarian judgment. Furthermore, whilst Koenigs *et al.* found that normal subjects rated personal scenarios as having significantly higher emotional salience than impersonal scenarios, they found no such significant difference between low and high-conflict scenarios. So their proposal that an affective deficit explains the VMPC patients' abnormal pattern of response to high-conflict scenarios is not clearly true. Similarly, it is unclear that this pattern of response is due to VMPC patients following "explicit social and moral norms"¹, since their choices in high conflict scenarios are contrary to familiar social norms to prevent harm.

In conclusion, to establish that a response pattern manifests a tendency to utilitarian moral judgement, the stimuli used need to be classified in terms of content and not by purely behavioural or emotional criteria as was done here and in other studies such as Greene *et al.*^{2,4,6}.

1. Koenigs M., Young, L., Adolphs R., Tranel, D., Cushman, F., Hauser, M., Damasio, A. (2007) Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature*, **446**, 908-911.
2. Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M. and Cohen, J. D. (2001) An fMRI investigation of emotional engagement in moral judgment. *Science*, **293**, 2105-2108.
3. Greene, J.D., & Haidt, J. (2002) How (and where) does moral judgment work? *Trends in Cognitive Sciences*, **6**, 517-523.
4. Greene, J.D. , Nystrom, L.E., Engell, A.D., Darley, J.M., and Cohen, J.D. (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron*, **44**, 389-400.
5. Singer, P. (2005) 'Ethics and Intuitions', *Journal of Ethics*, **9**: 331-352.
6. Ciaramelli, E., Muccioli, M., Làdavas, E. & di Pellegrino, G. (2007). Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cognitive and Affective Neuroscience*, **2**: 84-9.

Dr Guy Kahane
Deputy Director
Oxford Uehiro Centre for Practical Ethics
University of Oxford
Littlegate House
16/17 St Ebbes Street
Oxford
OX1 1PT
guy.kahane@philosophy.ox.ac.uk

Dr Nicholas Shackel
Lecturer in Philosophy
Department of Philosophy
ENCAP, University of Cardiff
Humanities Building
Colum Drive
Cardiff
CF10 3EU
+44 (0)2920 874025
shackeln@cardiff.ac.uk

and

Research Associate
Future of Humanity Institute
Faculty of Philosophy & James Martin 21st Century School
University of Oxford