Do abnormal responses show utilitarian bias?

Arising from: M. Koenigs *et al.* *Nature* 446, 908–911 (2007)

Neuroscience has recently turned to the study of utilitarian and non-utilitarian moral judgement. Koenigs *et al.*[1] examine the responses of normal subjects and those with ventromedial–prefrontal–cortex (VMPC) damage to moral scenarios drawn from functional magnetic resonance imaging studies by Greene *et al.*[2–4], and claim that patients with VMPC damage have an abnormally ''utilitarian'' pattern of moral judgement. It is crucial to the claims of Koenigs *et al.* that the scenarios of Greene *et al.* 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, who judge such choices to be appropriate, areas associated with cognition and cognitive conflict are activated as well. On the basis of a later study reporting similar results in responses to 'difficult' personal scenarios, Greene 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 ethicists[5].

Koenigs *et al.* draw on the battery of moral scenarios of Greene *et al.* 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 that 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 show 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. The distinction by Koenigs *et al.* 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 eight as pure cases of utilitarian versus 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 judgement. Furthermore, whereas 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], as 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 those of Greene *et al.*[2,4,6].

Guy Kahane[1] & Nicholas Shackel[2,3]

[1]Oxford Uehiro Centre for Practical Ethics, University of Oxford, Oxford OX1 1PT, UK.

[2]Department of Philosophy, ENCAP, University of Cardiff, Cardiff CF10 3EU, UK.

e-mail: shackeln@cardiff.ac.uk

[3]Future of Humanity Institute, Faculty of Philosophy & James Martin 21st Century School, University of Oxford, Oxford OX1 1PT, UK.

Received 29 August 2007; accepted 17 January 2008.

1. Koenigs, M. *et al.* Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* **446,** 908–911 (2007).
2. Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M. & Cohen, J. D. An fMRI investigation of emotional engagement in moral judgment. *Science* **293,** 2105–2108 (2001).
3. Greene, J. D. & Haidt, J. How (and where) does moral judgment work? *Trends Cogn. Sci.* **6,** 517–523 (2002).
4. Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M. & Cohen, J. D. The neural bases of cognitive conflict and control in moral judgment. *Neuron* **44,** 389–400 (2004).
5. Singer, P. Ethics and intuitions. *J. Ethics* **9,** 331–352 (2005).
6. Ciaramelli, E., Muccioli, M., Làdavas, E. & di Pellegrino, G. Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cogn. Affect. Neurosci.* **2,** 84–89 (2007).

doi:10.1038/nature06785

Koenigs et al. *reply*

Replying to: G. Kahane & N. Shackel *Nature* 452, doi:10.1038/06785 (2008)

Kahane and Shackel argue[1], on the basis of a re-classification of the moral scenarios used in our study[2], that our conclusion of a utilitarian bias among patients with ventromedial–prefrontal–cortex (VMPC) damage is unwarranted. Here we provide a re-analysis of our data based on precisely the classification scheme that Kahane and Shackel suggest. This re-analysis confirms our conclusion that damage to the VMPC results in an increase in utilitarian judgements.

Kahane and Shackel propose a classification scheme based solely on assessments of the scenario content. They suggest that utilitarian responses pertain only to those scenarios that pit ''consequences'' versus ''duty.'' We neither endorse nor disagree with this view; both their and our classification schemes are defensible.

In a re-analysis of our original data on the basis of the classification scheme suggested by Kahane and Shackel, we find that VMPC

patients generated the "utilitarian" judgement (as defined by Kahane and Shackel) in a substantially greater proportion than did either control group (71% by the VMPC group compared to 51% and 49% by the healthy and brain-damaged control groups, respectively; multinomial logistic regression, $P = 0.012$). Furthermore, among the 15 scenarios that present a utilitarian option, there was not one case where either control group endorsed a greater proportion of "utilitarian" responses than the VMPC group. We should note that this pattern of greater endorsement by the VMPC group was specific to the "consequence versus duty" scenarios: for the 9 "self-interest versus duty" moral scenarios in Kahane and Shackel's scheme, VMPC patients endorsed the proposed action in similar proportions to control groups (6% by the VMPC group compared to 2% and 10% by the healthy and brain-damaged control groups, respectively; $P = 0.31$). Likewise, in all 9 of the "self-interest versus duty" scenarios in Kahane and Shackel's scheme, at least one control group endorsed the proposed action in the same or greater proportion than did the VMPC group.

Kahane and Shackel also suggest that our results fail to demonstrate a causal role for emotion in moral judgements, because low- and high-conflict scenarios do not differ in emotional salience yet show differential effects of VMPC damage on moral judgements. Although the harms described in low- and high-conflict scenarios may be similarly emotionally salient, we reiterate that only in the high-conflict scenarios do these emotionally salient harms constitute morally ambiguous actions—in the low-conflict scenarios the emotionally aversive harms are quickly and unanimously condemned. In these scenarios, VMPC patients give normal responses, relying, we propose, on their capacity to use learned social rules, such as rules against harming others purely for self-interest.

This pattern of findings, together with VMPC patients' defects in processing social emotions, makes a causal role for emotion in moral judgement a plausible interpretation. This interpretation is consistent with studies showing that independent manipulations of emotion can influence moral judgement[3,4]. Furthermore, the main result from our original study (a selective effect of VMPC damage on moral judgement) has recently been replicated[5]. A final piece of data that is so far missing is concurrent monitoring of psychophysiological indices of emotion while subjects respond to moral scenarios, a technically challenging approach given the complex and temporally extended nature of the stimuli.

In summary, the re-analysis supports our original conclusion that VMPC patients are abnormally utilitarian in their moral judgement, regardless of how "utilitarian" is defined. Although we disagree with Kahane and Shackel about the conclusions of our original study, we certainly share the view that precise characterizations of distinct brands of moral judgement will prove fruitful in future studies of normal and pathological moral cognition[6–8].

M. Koenigs[1]†, L. Young[2], R. Adolphs[1,3], D. Tranel[1], F. Cushman[2], M. Hauser[2] & A. Damasio[1,4]

[1]Department of Neurology, University of Iowa Hospitals and Clinics, Iowa City, Iowa 52242, USA.
[2]Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, USA.
[3]Division of Humanities and Social Sciences and Division of Biology, California Institute of Technology, Pasadena, California 91125, USA.
[4]Brain and Creativity Institute and Dornsife Center for Cognitive Neuroimaging, University of Southern California, Los Angeles, California 90089, USA.
†Present address: National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892-1440, USA.
e-mail: radolphs@hss.caltech.edu

1. Kahane, G. & Shackel, N. Do abnormal responses show utilitarian bias? *Nature* **452,** 10.1038/06785 (2008).
2. Koenigs, M. *et al.* Damage to the prefrontal cortex increases utilitarian moral judgements. *Nature* **446,** 908–911 (2007).
3. Wheatley, T. & Haidt, J. Hypnotic disgust makes moral judgments more severe. *Psychol. Sci.* **16,** 780–784 (2005).
4. Valdesolo, P. & DeSteno, D. Manipulations of emotional context shape moral judgment. *Psychol. Sci.* **17,** 476–477 (2006).
5. Ciaramelli, E., Muccioli, M., Ladavas, E. & di Pellegrino, G. Selective deficit in personal moral judgment following damage to ventromedial prefrontal cortex. *Social Cogn. Affect. Neurosci.* **2,** 84–92 (2007).
6. Hauser, M. D. *Moral Minds: How Nature Designed a Universal Sense of Right and Wrong* (Harper Collins, New York, 2006).
7. Cushman, F. A., Young, L. & Hauser, M. D. The role of conscious reasoning and intuitions in moral judgment: testing three principles of harm. *Psychol. Sci.* **17,** 1082–1089 (2006).
8. Young, L., Cushman, F., Hauser, M. & Saxe, R. The neural basis of the interaction between theory of mind and moral judgment. *Proc. Natl Acad. Sci. USA* **104,** 8235–8240 (2007).

doi:10.1038/nature06804