ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 30 (2006) 1449 – 1457

The neural basis of implicit moral attitude—An IAT study using event-related fMRI

Qian Luo,[a] Marina Nakic,[a] Thalia Wheatley,[b] Rebecca Richell,[a] Alex Martin,[b] and R. James R. Blair[a,*]

[a]*Unit on Affective Cognitive Neuroscience, Mood and Anxiety Program, National Institute of Mental Health, Department of Health and Human Services, 15K North Drive, Room 206, MSC 2670, Bethesda, MD 20892, USA*
[b]*Laboratory of Brain and Cognition, National Institute of Mental Health, Bethesda, MD 20892, USA*

Received 27 July 2005; revised 1 November 2005; accepted 2 November 2005
Available online 18 January 2006

Recent models of morality have suggested the importance of affect-based automatic moral attitudes in moral reasoning. However, previous investigations of moral reasoning have frequently relied upon explicit measures that are susceptible to voluntary control. To investigate participant's automatic moral attitudes, we used a morality Implicit Association Test (IAT). Participants rated the legality of visually depicted legal and illegal behaviors of two different intensity levels (e.g., high intensity illegal = interpersonal violence; low intensity illegal = vandalism) both when the target concept (e.g., illegal) was behaviorally paired with an associated attribute (e.g., bad; congruent condition) or an unassociated attribute (e.g., good; incongruent condition). Behaviorally, an IAT effect was shown; RTs were faster in the congruent rather than incongruent conditions. At the neural level, implicit moral attitude, as indexed by increased BOLD response as a function of stimulus intensity, was associated with increased activation in the right amygdala and the ventromedial orbitofrontal cortex. In addition, performance on incongruent trials relative to congruent trials was associated with increased activity in the right ventrolateral prefrontal cortex (BA 47), left subgenual cingulate gyrus (BA 25), bilateral premotor cortex (BA 6) and the left caudate. The functional contributions of these regions in moral reasoning are discussed.
Published by Elsevier Inc.

Keywords: Moral attitude; IAT; Emotion; Amygdala; Ventrolateral PFC; Medial OFC

Introduction

Until relatively recently, most models of moral decision-making held a rationalist view (Kohlberg and Kramer, 1969; Piaget, 1932). Such models viewed moral reasoning as a conscious process; that is, it is attentional, effortful and controllable, and the reasoner is aware of what is going on (Bargh, 1994). However, recently, models stressing the role of emotion have become prevalent (Blair, 1995a,b; Greene and Haidt, 2002 review; Haidt, 2001; Kagan and Lamb, 1987; Moll et al., 2003 review). Part of the reason for this theoretical transition has been data collected from clinical populations. Thus, patients with damage to the ventromedial frontal cortex show no impairment for many aspects of reasoning yet are impaired in their emotional responses (e.g., Damasio et al., 1990), their moral emotions (Eslinger et al., 1992; Eslinger and Damasio, 1985) and their moral behavior (Anderson et al., 1999; Blair and Cipolotti, 2000; Damasio, 1999; Eslinger and Damasio, 1985; Grafman et al., 1996). Similarly, individuals with psychopathy show no impairment in many aspects of non-emotional reasoning (see Blair, 2004). However, they are impaired in specific forms of emotional responding (Blair et al., 2001; Lykken, 1957), their moral emotions (Blair, 1995a,b; Hare, 1991), their moral reasoning (Blair, 1995a,b; Gray et al., 2003) and their moral behavior (Hare, 1991). By understanding the neuro-cognitive systems involved in moral reasoning, we may increase our understanding of these clinical conditions. To further this goal, we investigated the neural systems underlying moral reasoning performed in the context of a moral Implicit Associations Task.

Several recent studies have investigated the neural systems involved in moral reasoning (e.g., Greene et al., 2001, 2004; Heekeren et al., 2003, 2005; Moll et al., 2001, 2002a,b). These studies have revealed the importance of medial orbitofrontal cortex (Greene et al., 2001, 2004; Heekeren et al., 2003, 2005; Moll et al., 2001, 2002a,b), the cingulate gyrus (Greene et al., 2001; Moll et al., 2001, 2002a,b), superior temporal sulcus (Heekeren et al., 2003, 2005; Moll et al., 2002a,b) and the amygdala (Greene et al., 2004; Heekeren et al., 2005; Moll et al., 2002a,b). However, the functional contributions of these regions remain relatively unclear.

The previous moral reasoning work has used varying methodologies such as making moral decisions based on text descriptions of ethical dilemmas (Greene et al., 2001), passive viewing pictures of moral violations (Moll et al., 2002a), judging sentence

* Corresponding author. Fax: +1 301 451 5783.
E-mail address: James.Blair@mail.nih.gov (R.J.R. Blair).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter. Published by Elsevier Inc.
doi:10.1016/j.neuroimage.2005.11.005

descriptions of behaviors as moral or immoral (Moll et al., 2002b; Heekeren et al., 2005) and making moral decisions (morally appropriate or not) versus semantic decisions (semantically correct or not) on sentences (Heekeren et al., 2003). One feature that these methodologies have in common is that they rely on explicit processing; the participant is asked to make an explicit judgment of a behavior (e.g., Greene et al., 2001, 2004; Heekeren et al., 2003, 2005; Moll et al., 2001). However, such measures are susceptible to voluntary control and allow a participant the ability to conceal their genuine attitudes. Moreover, recent work on moral reasoning has stressed its "automatic" nature (Greene and Haidt, 2002; Haidt, 2001).

One methodology that can be considered to assess an individual's automatic and implicit attitudes towards social stimuli is the Implicit Association Test (IAT; Greewald et al., 1998). This test measures the extent to which two target concepts (e.g., flower and insect) are associated with two attributes (e.g., good and bad). In contrast to verbal measures of self-report, the IAT relies on differentials in reaction times to index an individual's automatic attitudes. When the target concept (e.g., flower) is paired with an associated attribute (e.g., good), the participant's reaction times are faster than when it is paired with an attribute to which it is not associated (e.g., bad). The IAT can therefore be used to identify an individual's implicit attitudes, for example, to out-groups, regardless of the individual's wish to hide these attitudes (e.g., Greewald et al., 1998; Greenwald and Farnham, 2000). Recently, the IAT has been adapted to assess an individual's automatic attitudes to moral and immoral actions, even in individuals with psychopathy who might wish to conceal their attitude to these actions (Gray et al., 2003). Such studies reveal a reduced automatic "bad" attitude towards immoral actions in individuals with psychopathy relative to comparison populations.

Very little work has investigated the neural correlates of performance on the IAT. Phelps et al. (2000) found that the strength of an "automatic" amygdala response to racial out-groups predicted the level of IAT effect for a race-based task. However, this study did not examine neural correlates of performing the IAT task itself. Chee et al. (2000) examined the neural correlates of an IAT task involving the assessment of the association of two object categories (flower and insect) with the valenced attribute categories of "pleasant" and "unpleasant". Chee et al. (2000) reported that performance on "incongruent" trials (trials where the same response is made for stimuli associated with differently valenced attributes; e.g., "flower" and "unpleasant") was associated with significantly greater activity in ventrolateral (BA 47), dorsolateral (BA 9, 44) prefrontal cortex and anterior cingulate (BA 32). However, there has been no fMRI investigation of a morality IAT task.

In the current study, we aimed to determine regions involved in the performance of the morality IAT task. We predicted on the basis of prior fMRI data investigating moral reasoning (Greene et al., 2004; Heekeren et al., 2005; Moll et al., 2002a,b) that the individual's automatic moral response (as indexed by an increased response to high intensity stimuli [interpersonal violence] relative to low intensity stimuli [vandalism]) would recruit the amygdala, superior temporal sulcus and medial orbital frontal cortex. In addition, we predicted on the basis of Chee's earlier study (Chee et al., 2000) that the IAT effect (as indexed by an increased response to incongruent trials rather than congruent trials) would be related to increased activity in ventrolateral prefrontal cortex and anterior cingulate.

Materials and methods

Participants

Twenty healthy volunteers, 9 males and 11 females, between the ages of 20 to 36 participated in this study. All gave written informed consent to participate in the study, which was approved by the National Institute of Mental Health Institutional Review Board.

The morality implicit association task and experimental procedure

The stimuli consisted of 48 color photographic stimuli primarily selected from the International Affective Picture System (IAPS; Lang and Greenwald, 1985); several of the low intensity illegal items were taken from additional sources. As some of the low intensity illegal stimuli were taken from additional sources, we examined participant's ratings of their pleasantness and autonomic responsiveness to all test stimuli (Blair et al., submitted for publication). These are reported together with the IAPS averages in Table 1. The stimuli consisted of images of: 8 highly arousing illegal behaviors (for example, these stimuli involved interpersonal violence; e.g., guns/knives used in attacks/as threats); 8 less arousing illegal behaviors (for example, these stimuli involved property damage but did not involve interpersonal violence); 8 highly arousing legal behaviors (e.g., skydiving); and 8 less arousing legal behaviors (e.g., playing the guitar). Legal behaviors were chosen rather than moral behaviors because of the difficulty of obtaining sufficient understandable stimuli (for example, depictions of charity work are difficult to parse rapidly). However, work suggests that judgments about what is legal are based on judgments about what they consider to be moral (Greene and Cohen, 2004; Helwig and Jasiobedzka, 2001). In addition, there were pictures of 8 negative animals (e.g., snake) and 8 positive animals (e.g., puppies). Following the previous literature (e.g., Greewald et al., 1998), the fMRI experiment comprised of two series of 5 phases. Within each of the two series, a different set of stimuli was used; for example, 4 high arousing illegal stimuli were used during the first series of five phases and a different 4 used during the second series of five phases. Only data recorded during phases 3 and 5 in each of these two series were recorded.

The ten phases were:

Phase 1 the subject categorized behaviors as legal (left hand) or illegal (right hand);

Phase 2 the subject judged animals as good (left hand) or bad (right hand).

Table 1
Ratings and autonomic responses to the stimuli presented

	IAPS valence	IAPS arousal	Pleasantness	Arousal
High illegal	2.48	6.52	−4.45	0.0110 microsiemens
Low illegal	–	–	−1.33	0.0049 microsiemens
High legal	6.98	6.06	1.86	0.0106 microsiemens
Low legal	5.2	3.44	0.77	0.0055 microsiemens
Bad animal	3.97	5.59	–	–
Good animal	7.49	4.17	–	–

IAPS: International Affective Picture Scale (Lang et al., 1995).