

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,*}

^aUnit 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

^bLaboratory 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).

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; Greenwald 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., Greenwald 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., Greenwald 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).

Phase 3 with behaviors and animals randomly presented, the subject judged if pictures were legal behaviors/good animals (left hand) or illegal behaviors/bad animals (right hand). This phase involved *congruent* trials; the target concepts (legal/illegal) were paired with their associated attributes (good/bad).

Phase 4 the subject categorized behaviors as legal (right hand) or illegal (left hand), opposite to that in Phase 1.

Phase 5 with behaviors and animals randomly presented, the subject judged if pictures were illegal behaviors/good animals (left hand) or legal behaviors/bad animals (right hand). This phase involved *incongruent* trials; the target concepts were not paired with their associated attributes (that is, illegal was paired with good).

The second series of 5 phases was identical to the first except that Phase 1 and Phase 4 swapped position. Thus:

Phase 6 the subject categorized behaviors as legal (right hand) or illegal (left hand), opposite to that in Phase 1 but identical to phase 4.

Phase 7 the subject judged animals as good (left hand) or bad (right hand).

Phase 8 the subject judged if pictures were illegal behaviors/good animals (left hand) or legal behaviors/bad animals (right hand); i.e., *incongruent* trials.

Phase 9 the subject categorized behaviors as legal (left hand) or illegal (right hand), opposite to that in Phases 4 and 6 but identical to Phase 1.

Phase 10 the subject judged if pictures were legal behaviors/good animals (left hand) or illegal behaviors/bad animals (right hand); i.e., *congruent* trials.

In short, the participants experienced a counterbalanced ABBA (congruent incongruent incongruent congruent)/BAAB (incongruent congruent congruent incongruent) design; half of the participants experienced congruent incongruent incongruent congruent and the other half incongruent congruent congruent incongruent phases. Fixation point trials served as the baseline.

Each trial lasted 2 s, one TR. Each trial consisted of a presentation of a picture (or fixation) for 1000 ms followed by a 1000 ms blank screen. Each of the 4 scanning sessions lasted for 8 min and 10 s and covered one congruent or one incongruent phase.

Each session involved 120 events and 120 fixations. Each event involved one of the 4 high-illegal, 4 high-legal, 4 low-illegal, 4 low-legal behavior pictures, 4 good and 4 bad animal pictures or a fixation point trial. Each of the pictures was seen five times in each of the congruent/incongruent phases.

MRI data acquisition

The study was performed on a 1.5 T GE scanner. Functional scans were obtained by using a single-shot T2*-weighted gradient-echo planar imaging (EPI) sequence (29 contiguous axial slices, slice thickness = 4 mm, in plane resolution = 3.75×3.75 , TR/TE/ θ = 2000 ms/30 ms/90, FOV = 240×240 mm², matrix = 64×64). A series of 240 images were acquired during each of the 4 functional runs. The high-resolution anatomical images were acquired using a T1-weighted, three-dimensional, Spoiled GRASS imaging (spgr) sequence ($1 \times 1 \times 1.5$ mm³).

Data processing

The AFNI software package (<http://afni.nimh.nih.gov/afni>) was used for image data processing. The images of the first five time points in each run were discarded. The 4 times series were motion-corrected and reregistered and spatially smoothed with Isotropic Gaussian blur with FWHM = 3 mm. They were then normalized to signal percent change. This was done by dividing the signal intensity of a voxel at each time point by the mean signal intensity of that voxel for each run and multiplying the result by 100. They were then concatenated. The hemodynamic response function (HRF) and multivariate statistics corresponding to each condition (high-illegal congruent, low-illegal congruent, high-legal congruent, low-legal congruent, high-illegal incongruent, low-illegal incongruent, high-legal incongruent, low-legal incongruent, 'good' animals and 'bad' animals, fixation point trials) were obtained by deconvolving the input for each from the concatenated time series using a least squares procedure within a General Linear Model. Individual images were spatially normalized to the Talairach brain atlas before group analysis.

Individual images were spatially normalized to the standard coordinate space of Talairach and Tournoux (Talairach and Tournoux, 1988) before group analysis. The group analysis was then performed using a 2 (IAT effect: incongruent/congruent) \times 2 (Stimulus: illegal/legal) \times 2 (Intensity: high/low) ANOVA. The threshold was set at $P < 0.0001$ (corrected at 0.01 for multiple comparison) for the main effects. Threshold correction was done by using the AlphaSim program in AFNI, which applies Monte Carlo simulation to calculate the probability of false positive detection by taking into consideration both the individual voxel probability thresholding and cluster size. See Table 2 for activation results.

Table 2
Regions that showed a significant main effect in voxel-wise analysis

Structure	L/R	BA	x	y	z	t
<i>Incongruent > congruent, P < 0.0001 (corrected at 0.01)</i>						
Ventrolateral prefrontal cortex	R	47	44	5	-6	5.31
Subgenual gyrus	L	25	-3	12	-15	5.07
Premotor cortex	L	6	-21	-1	60	5.09
Premotor cortex	R	6	33	0	43	5.30
Caudate	L		-16	11	11	4.92
Insula	L	13	-32	-14	19	5.27
Precuneus	L	31	-9	-63	26	5.62
Lentiform nucleus	R		18	-2	6	4.62
Cingulate gyrus	L	24	-4	7	29	5.20*
<i>High > low, P < 0.01</i>						
Amygdala	R		22	-6	-8	2.92
Medial orbital frontal cortex	L	10	-6	55	-5	2.83
<i>Illegal > legal, P < 0.0001 (corrected at 0.01)</i>						
Medial frontal gyrus	R	6	1	-3	60	4.72
Superior temporal sulcus	L	41	-49	-29	14	4.69
Superior temporal sulcus	R	41	42	-37	15	4.52
Superior temporal sulcus	L	22	-48	-17	9	4.76
Precentral gyrus	L	4	-34	-27	61	5.12
Precentral gyrus	L	4	-40	-17	57	4.89
Precentral gyrus	L	6	-30	-8	55	4.45
Postcentral gyrus	L	3	-35	-24	45	4.51
Amygdala	R		21	-7	-8	3.44*

* Indicates areas that did not survive correction for multiple comparisons.

Results

Behavioral data

Fig. 1 depicts the participant's RTs by stimulus condition and congruence. The participants made few errors (1.62%), and RTs for these trials were not included in subsequent analyses. A 2 (Congruence: congruent vs. incongruent) \times 2 (Stimulus intensity: High vs. Low) \times 2 (Legality: Illegal vs. Legal) ANOVA was conducted on the data. This revealed significant main effects for congruence, stimulus intensity and legality. The subjects responded significantly faster in the congruent than in the incongruent condition ($M_{\text{incongruent}} = 771.41 \pm 83.50$ ms, $M_{\text{congruent}} = 714.32 \pm 67.91$ ms, $F(1,19) = 51.804$, $P < 0.0001$). In short, the participants showed an IAT effect. In addition, RTs were shorter for the higher intensity than the lower intensity stimuli ($M_{\text{high}} = 736.90 \pm 66.46$ ms, $M_{\text{low}} = 748.82 \pm 77.12$ ms, $F(1,19) = 12.286$, $P < 0.002$) and illegal rather than legal stimuli ($M_{\text{illegal}} = 735.63 \pm 68.03$ ms, $M_{\text{legal}} = 750.10 \pm 83.38$ ms, $F(1,19) = 4.109$, $P < 0.057$). There were no significant interactions.

fMRI data

The ANOVA on the BOLD response data yielded significant main effects, but no significant interactions were found. The results of each main effect are described in turn.

The IAT effect: incongruent vs. congruent trials

The first main effect identified regions which showed a differential BOLD response to incongruent trials relative to congruent trials. These included right ventrolateral prefrontal cortex (BA 47), left subgenual gyrus (BA 25), bilateral premotor cortex (BA 6), left caudate, insular and precuneus and the right lentiform nucleus, all of which showed significantly greater activity to incongruent relative to congruent trials (see Table 1). There was also evidence of significantly greater cingulate (BA 24) activity ($P < 0.0001$, although this did not survive correction for multiple comparisons). From this contrast and our a priori hypotheses, functionally defined ROIs were identified in right ventrolateral prefrontal cortex (BA 47), left medial frontal gyrus

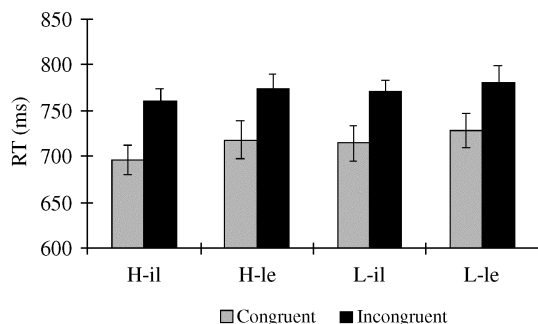


Fig. 1. Mean reaction times (RTs) across conditions. Legends: key to this figure: H-il = high illegal, H-le = high legal, L-il = low illegal, L-le = low legal. Congruent = trials where the same response was made to illegal stimuli and 'bad' animals and a different same response was made to legal and 'good' animals. Incongruent = trials where the same response was made to illegal stimuli and 'good' animals and a different same response was made to legal and 'bad' animals.

(BA 25), the left caudate and the left cingulate gyrus (BA 24; see Figs. 2A–D).

Intensity: high vs. low

The second main effect identified regions which showed a differential response to high relative to low intensity legal and illegal stimuli. This contrast did not reveal any significant differences at the threshold adopted for the congruency effect ($P < 0.0001$ corrected for multiple comparisons at $P < 0.01$). However, given our a priori hypotheses, we used anatomically specified ROIs of bilateral amygdala and Brodmann area 10 and examined whether there were regions within these ROIs that showed greater BOLD responses to high relative to low intensity legal and illegal stimuli. This revealed significant activations in both regions ($P < 0.01$ uncorrected; see Figs. 2E and F).

Stimulus: illegal vs. legal trials

The third main effect identified regions which showed a differential response to illegal relative to legal stimuli. These included right medial frontal gyrus (BA 6), left precentral gyrus (BA4, 6) and postcentral gyrus (BA 3) and bilateral superior temporal gyrus (BA 22, 41), all of which showed significantly greater activity to illegal relative to legal trials (see Table 1). There was also evidence of significantly greater amygdala activity ($P < 0.005$, although this did not survive correction for multiple comparisons).

Discussion

The goal of the current study was to determine regions involved in the performance of the morality IAT task. Behaviorally, and in line with previous work using other types of IAT paradigm (e.g., Chee et al., 2000; Gray et al., 2003; Greewald et al., 1998), participants were significantly slower on incongruent (when categories of different valence shared the same key: e.g., moral transgressions and good animals) than on congruent trials (when categories of the same valence shared the same key: e.g., moral transgressions and bad animals); that is, the participants showed the IAT effect. This is consistent with our suggestion that healthy individuals have an "automatic" association between, in particular, illegal/immoral behaviors and negative valence (cf. Blair, 1995a,b; Haidt, 2001; Nichols, 2002). With respect to the fMRI data, we found three determinants of the BOLD response: the congruence/incongruence of the trial, stimulus intensity and stimulus legality. These will be discussed in turn.

The neural correlates of the IAT effect

With respect to the BOLD response, participants showed significantly greater activity within anterior cingulate gyrus (BA 24), the subgenual portions of the anterior cingulate gyrus (BA 25), right ventrolateral frontal gyrus (BA 47), bilateral premotor cortex (BA 6) and caudate in response to incongruent relative to congruent trials. In the only previous study examining the neural correlates of the IAT effect, Chee et al. (2000) similarly reported that incongruent, relative to congruent, trials were associated with significantly greater activity in anterior cingulate and ventrolateral frontal cortex.

The IAT task (the increase in RT during incongruent relative to congruent trials) shares features with both the go/no-go and

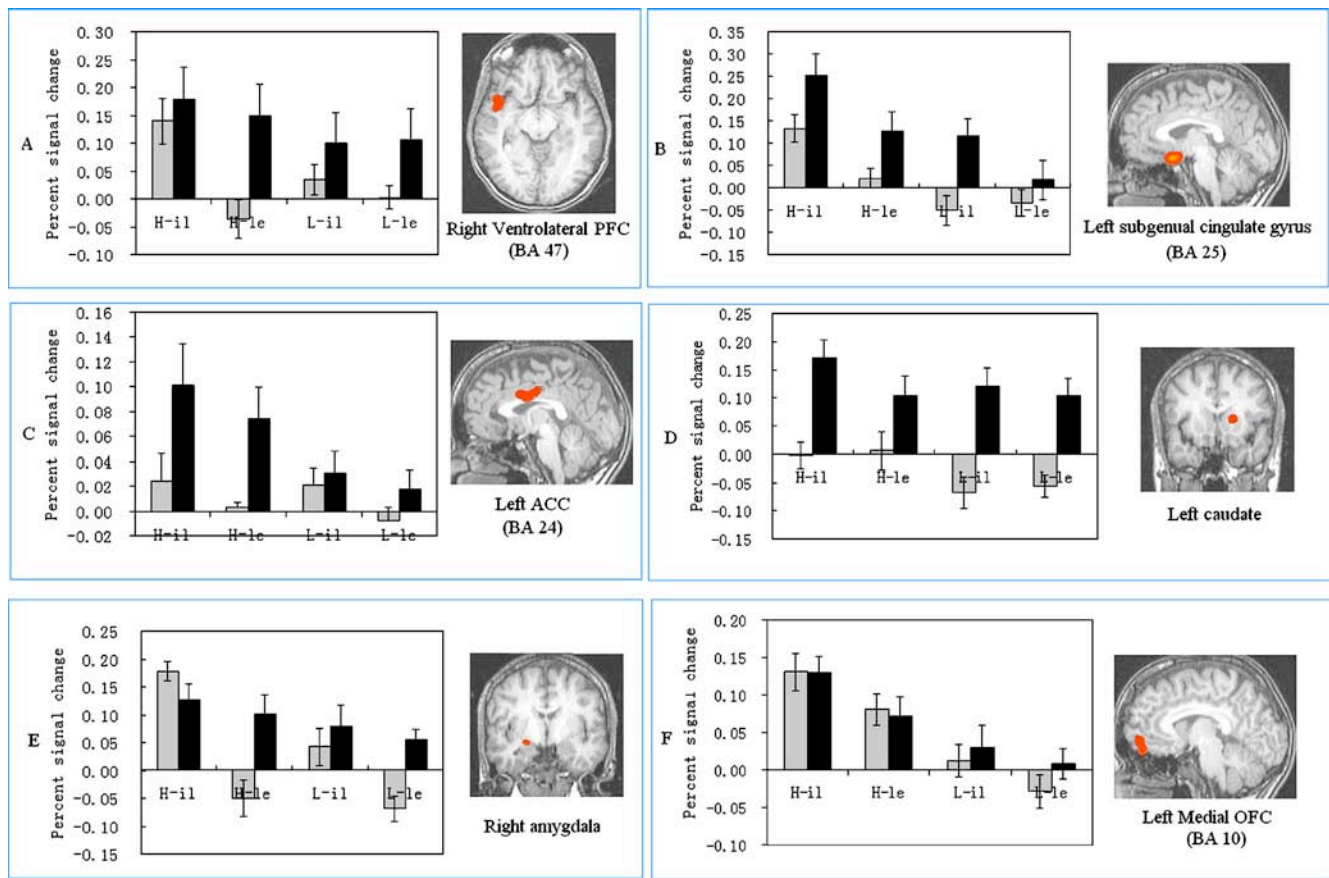


Fig. 2. Signal percent change and activation maps for brain areas. Legends: A–D: areas activated by incongruent trials relative to congruent trials. Activations, and associated percentage signal changes from baseline, are shown for: (A) Right ventrolateral prefrontal cortex (IFG); (B) left subgenual cingulate; (C) left anterior cingulate (ACC); and (D) left caudate. (E–F) Areas activated by high intensity relative to low intensity trials. Activations, and associated percentage signal changes from baseline, are shown for: (E) right amygdala; and (F) left medial orbital frontal cortex (OFC). Key to this figure: H-il = high illegal, H-le = high legal, L-il = low illegal, L-le = low legal. Congruent trials are depicted in grey. Incongruent trials are depicted in black.

response reversal tasks; all three tasks require the activation of an alternative motor response to one strongly activated by the stimulus conditions. The IAT task requires the participant to over-rule a response tendency that is more consistent with the valence of the stimulus to be responded to. The go/no-go task requires the participant to withhold from responding to a specific infrequent stimulus while responding to other more common stimuli (Casey et al., 2001). Response reversal requires the participant to change their response to a pair of stimuli following changes in stimulus–reinforcement contingencies (Cools et al., 2002). Interestingly, responses during no-go trials in go/no-go tasks and to reversal errors (triggering response change) in response reversal tasks implicate similar regions to those recruited during incongruent trials in the present study. In short, responses during no-go trials in go/no-go tasks and to reversal errors are associated with activity in anterior cingulate (Kringelbach and Rolls, 2003; Nagahama et al., 2001; O’Doherty et al., 2003), ventrolateral prefrontal cortex (Casey et al., 2001; Cools et al., 2002; O’Doherty et al., 2003) and caudate (Kringelbach and Rolls, 2003; O’Doherty et al., 2003).

Anterior cingulate has been implicated in monitoring and controlling goal-directed behavior (Bush et al., 2000; Kerns et al., 2004; Schall et al., 2002) and particularly conflict monitoring (Botvinick et al., 1999; Botvinick et al., 2004; Carter et al., 1998;

Cohen et al., 2000; MacDonald et al., 2000). Anterior cingulate activation is frequently seen in the presence of dorsolateral prefrontal cortex activation in the context of attentional tasks such as the Stroop task (Bush et al., 1998; Leung et al., 2000; MacLeod and MacDonald, 2000; Pardo et al., 1990). Premotor cortex (BA 6) and the caudate are both importantly involved in mediating motor responding (Purves et al., 2001; Mink, 1996). Both regions have frequently been observed to show greater activation in conditions of response conflict (Bush et al., 1998; Bush et al., 1999; Kerns et al., 2004; Casey et al., 2001; 2002; Durston et al., 2002, 2003; Menon et al., 2001; Mink, 1996; Nunez et al., 2005; Peterson et al., 2002) including responses in the go/no-go tasks (e.g., Durston et al., 2002; Menon et al., 2001).

With respect to attentional tasks such as the Stroop task, the suggestion has been made that anterior cingulate monitors conflict. It has been considered to potentially augment the representation of *stimulus features* by dorsolateral prefrontal cortex leading to increased cognitive control of stimuli represented in temporal cortex (Garavan et al., 2002; MacDonald et al., 2000; Ruff et al., 2001). As noted above, there are frequent common activations of anterior cingulate and ventrolateral frontal cortex in the context of go/no-go and response reversal tasks. Moreover, both regions were also seen in the only preceding study of the IAT effect (Chee et al., 2000). We suggest that anterior cingulate following motor response

conflict may also augment the representation of object/motor features within ventrolateral frontal cortex that allow control over motor responding mediated by the caudate.

The neural correlates of automatic moral attitude

We observed significantly greater activity within the right amygdala and the left medial orbitofrontal cortex during high intensity (e.g., interpersonal violence) relative to low intensity (vandalism) trials. It should be noted that the high intensity illegal stimuli involved interpersonal violence, while the low intensity illegal stimuli involved property damage. Under some conceptualizations of care-based morality, direct interpersonal harm is a special case (Greene et al., 2001, 2004). Under others (Blair, 1995a,b), a transgression is considered immoral to the extent to which it has been associated with victims; highly intense moral transgressions are those closely associated with victims (in the current stimuli, the victims were depicted) while for less intense moral transgressions the association is weaker (in the current study, property damage is likely to sadden the owners of the property, but these “victims” were not depicted in the stimuli). The current study was not designed to distinguish between these positions. However, it is possible that the increase in amygdala/medial orbital frontal cortex observed here to the more intense illegal stimuli was related to the direct interpersonal harm depicted rather than the intensity of the stimuli per se.

The amygdala has long been associated with emotional processing (LeDoux, 1998) and, in neuroimaging work, is consistently activated by emotional relative to neutral stimuli (see Phan et al., for a review). With respect to morality, previous work has shown significant amygdala activation when viewing scenes of moral violations (Moll et al., 2002a), judging sentence descriptions of behaviors as moral or immoral (Moll et al., 2002b; Heekeren et al., 2005) and making moral decisions based on text descriptions of ethical dilemmas (Greene et al., 2004).

Phelps et al. (2000) reported that the strength of healthy participants’ IAT effect for a race-based IAT task was positively associated with the participant’s amygdala response to images of individuals from the racial out-group. This suggests a role for the amygdala in contributing the valence information necessary to generate the IAT effect. However, two caveats should be made here: first, a single case study of a patient with an acquired amygdala lesion did suggest intact race bias as measured by the IAT despite the amygdala lesion (Phelps et al., 2003); and second, there are many forms of IAT effect, e.g., gender-based and even math-based (Nosek and Banaji, 2002; Greenwald and Banaji, 1995). While a threat-based account of out racial group/anti-mathematics bias may seem plausible, it is less clear that a threat-based account of out gender bias IAT effects is appropriate. Of course, it is possible that gender bias IAT effects are based on same valence facilitation (in gender associated with good) rather than different valence interference (in gender associated with bad) as well as same valence facilitation. If so, this predicts that incongruent trials in a gender-based IAT should not be associated with the anterior cingulate, ventrolateral frontal gyrus and caudate response seen here.

In the current study, we observed increased amygdala activity to the more intense stimuli (illegal: kidnapping; legal: figure-skating) than to the less intense stimuli (illegal: vandalism; legal: people chatting). In our previous behavioral work, we have found an increased IAT effect for more intense moral transgressions than less

intense moral transgressions (Blair et al., submitted for publication). Moreover, individuals with psychopathy, who show impairment on a range of functions reliant on the amygdala (see Blair, 2004), show a reduced IAT effect on the current, or similar, tasks (Blair et al., submitted for publication; Gray et al., 2003). In short, we believe the current and previous data suggest that the amygdala contributes the valence information necessary (the “automatic moral attitude”) to generate the IAT effect.

The amygdala is considered to be crucially involved in the formation of stimulus–reinforcement associations (Everitt et al., 2003; LeDoux, 1998). In line with this, we believe the amygdala plays a role in morality by allowing the association of representations of transgressions (interpersonal violence) with the aversive stimulus of the victim’s fear/sadness (Blair, 1995a,b; Blair, 2001). We believe that a component, at the neural level, of the individual’s “automatic moral attitude” to a moral transgression involves the activation of the amygdala. This is elicited by the conditioned stimulus that is the individual’s representation of the moral transgression.

Previous work has shown that medial regions of orbitofrontal cortex show greater responding to emotional rather than neutral stimuli (Dolcos et al., 2004; Liberzon et al., 2000; Ochsner et al., 2002). Several researchers have suggested that this medial orbitofrontal cortex response is particularly involved in processing rewarding stimuli (Elliott et al., 2000; O’Doherty et al., 2001; Rolls, 2000). Some data have been consistent with this position (Dolcos et al., 2004; Elliott et al., 2000; O’Doherty et al., 2001; Rolls, 2000). Alternatively, medial orbitofrontal cortex has been considered to serve a role in decision-making/response selection (Bechara et al., 2000), perhaps using expected reinforcement (both positively and negatively valenced) information to guide response/stimulus choice (Blair, 2004). Recent studies of morality using different 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 descriptions of behaviors as moral or immoral (Heekeren et al., 2005; Moll et al., 2002b) and making moral decisions (morally appropriate or not) versus semantic decisions (semantically correct or not) on sentences (Heekeren et al., 2003) have all implicated medial regions of orbitofrontal cortex. The activations reported in these studies correspond very closely to the medial orbitofrontal cortex activation seen in the current study in response to higher salience rather than lower salience stimuli. Importantly, in the current study, as well as in the previous morality work, this medial orbitofrontal cortical activation was seen to both transgressions (illegal behaviors) as well as positively valenced legal behaviors. This is consistent with suggestions that this region is involved in decision-making/response selection as a function of expected reinforcement information (Bechara et al., 2000; Blair, 2004).

In short, we would propose that an individual’s automatic moral attitude to an event involves an integrated neural response involving both the amygdala and medial orbitofrontal cortex that is proportional to the emotive strength (due to previous learning) of the stimulus. This “emotive strength” (the reward/punishment value associated with the stimulus) provides information regarding the “badness” of the action and plays a role in generating the IAT effect, if stimuli associated with opposite valences are associated with the same response. Consistent with this suggestion, lesions of medial orbitofrontal cortex have been reported to abolish the IAT effect (Milne and Grafman, 2001). However, it is important to note that, while “emotive strength” provides information regarding the

“badness” of the action, a full judgment of an action’s immorality is more than an automatic moral attitude of the action’s “badness” (cf. Nichols, 2002). The suffering caused by natural disasters is “bad,” and images of such suffering would likely lead to amygdala/medial orbitofrontal cortex activity, but the natural disasters themselves are not considered immoral. To consider an action to be immoral, we usually ascribe intent to the actor (cf. Nichols, 2002). Moreover, it is important to note that there are additional types of moral reasoning/concepts that extend beyond the automatic moral attitude to harm considered in the present paper (e.g., notions of fairness, respect, duty, purity, loyalty, honor).

Illegal and legal stimuli

While we had no a priori predictions regarding a differential neural response to the illegal and legal stimuli (the stimuli sets had been matched for verbal and autonomic ratings of arousal), we found greater BOLD responses to illegal relative to legal stimuli in motor and premotor cortex, superior temporal cortex and the amygdala. In addition, behaviorally, we found that RTs were faster to illegal relative to legal stimuli. This is consistent with previous literature suggesting that events that are negatively valenced can have a greater impact on the individual than positively valenced events of the same type (Baumeister et al., 2001; Rozin and Royzman, 2001). However, it is worth noting that our illegal items were immoral (e.g., interpersonal violence/property damage), while our legal items were not moral per se but simply legal (e.g., skydiving/typing). Several previous studies have implicated superior temporal sulcus in moral reasoning (Heekeren et al., 2003, 2005; Moll et al., 2002a,b). Superior temporal sulcus is thought to be involved in processing intentionality (e.g., Frith and Frith, 1999; Pelphrey et al., 2004). Intention information is, of course, crucial for moral reasoning (Piaget, 1932; Kohlberg, 1969). However, the role of intention information in moral reasoning was not the focus of the current study. Further work is needed to delineate the role of STS in moral reasoning.

Conclusions and implications

The present study investigated the neural basis of implicit moral attitudes using the IAT paradigm. We suggest that activity within the amygdala and medial orbitofrontal cortex in our own and previous studies of moral decision-making reflects the generation of a moral attitude; i.e., the activation of a stimulus–reinforcement association [representation of moral transgression-negative valence association] and the representation of the negative valence expectancy, following amygdala input, by medial orbitofrontal cortex. Resolving the motor conflict induced by differentially valenced objects associated with the same response option leads to increased BOLD responses within anterior cingulate, ventrolateral frontal cortex and caudate. We suggest that anterior cingulate responds to this motor response conflict by augmenting the representation of object/motor features within ventrolateral frontal cortex that allow control over *motor* responding mediated by the caudate. With respect to neuropsychiatric and neurological populations, we should expect that disruption of the ability to form stimulus–reinforcement associations, as is seen in individuals with psychopathy (Blair, 2004), or the ability represent valence expectancies, as is seen after ventromedial frontal cortical lesions, should disrupt morality IAT effects (see Gray et al., 2003; Milne

and Grafman, 2001). There should be no response conflict induced by differentially valenced objects because either the stimulus–reinforcement associations have not been formed or the medial orbitofrontal cortex can no longer represent them. Interestingly, though lesions of anterior cingulate or ventrolateral frontal cortex should augment the IAT effect, the system will be less able to augment the activity of the appropriate motor response. Future work will determine the validity of this last prediction.

References

- Anderson, S.W., Bechara, A., Damasio, H., Tranel, D., Damasio, A.R., 1999. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nat. Neurosci.* 2, 1032–1037.
- Bargh, J.A., 1994. The four horsemen of automaticity: awareness, intention, efficiency, and control in social cognition. In: Wyer, R.S., Srull, T.K. (2nd ed.). *Handbook of Social Cognition* vol. 1. Erlbaum, Hillsdale, NJ, pp. 1–40.
- Baumeister, R.F., Bratslavsky, E., Finkenauer, C., Vohs, K.D., 2001. Bad is stronger than good. *Rev. Gen. Psychol.* 5, 323–370.
- Bechara, A., Damasio, H., Damasio, A.R., 2000. Emotion, decision making and the orbitofrontal cortex. *Cereb. Cortex* 10, 295–307.
- Blair, R.J.R., 1995a. A cognitive developmental approach to morality: investigating the psychopath. *Cognition* 57, 1–29.
- Blair, R.J.R., 1995b. A cognitive developmental approach to morality: investigating the psychopath. *Cognition* 57, 1–29.
- Blair, R.J., 2001. Neurocognitive models of aggression, the antisocial personality disorders, and psychopathy. *J. Neurol. Neurosurg. Psychiatry* 71, 727–731.
- Blair, R.J., 2004. The roles of orbital frontal cortex in the modulation of antisocial behavior. *Brain and Cognition* 55, 198–208.
- Blair, R.J.R., Cipolletti, L., 2000. Impaired social response reversal: a case of “acquired sociopathy”. *Brain* 123, 1122–1141.
- Blair, R.J.R., Monson, J., Frederickson, N., 2001. Moral reasoning and conduct problems in children with emotional and behavioural difficulties. *Pers. Individ. Differ.* 31, 799–811.
- Botvinick, M., Nystrom, L., Fissell, K., Carter, C., Cohen, J., 1999. Conflict monitoring vs. selection-for-action in anterior cingulate cortex. *Nature* 402, 179–181.
- Botvinick, M., Cohen, J., Carter, C., 2004. Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn. Sci.* 12, 539–546.
- Bush, G., Whalen, P.J., Rosen, B.R., Jenike, M.A., McInerney, S.C., Rauch, S.L., 1998. The counting Stroop: an interference task specialized for functional neuroimaging-validation study with functional MRI. *Hum. Brain Mapp.* 6, 270–282.
- Bush, G., Frazier, J.A., Rauch, S.L., Whalen, P.J., Jenike, M.A., Rosen, B.R., Biedermann, J., 1999. Anterior cingulate cortex dysfunction in attention-deficit/hyperactivity disorder revealed by fMRI and counting Stroop. *Biol. Psychiatry* 45, 1542–1552.
- Bush, G., Luu, P., Posner, M.I., 2000. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn. Sci.* 4, 215–222.
- Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., Cohen, J.D., 1998. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280, 747–749.
- Casey, B.J., Forman, S.D., Franzen, P., Berkowitz, A., Braver, T.S., Nystrom, L.E., Thomas, K.M., Noll, D.C., 2001. Sensitivity of prefrontal cortex to changes in target probability: a functional MRI study. *Hum. Brain Mapp.* 13, 26–33.
- Casey, B.J., Thomas, K.M., Davidson, M.C., Kunz, K., Franzen, P.L., 2002. Dissociating striatal and hippocampal function developmentally with a stimulus–response compatibility task. *J. Neurosci.* 22, 8647–8652.
- Chee, M.W.L., Sriram, N., Soon, C.S., Lee, K.M., 2000. Dorsolateral prefrontal cortex and the implicit association of concepts and attributes. *NeuroReport* 11, 135–140.

- Cohen, J.D., Botvinick, M., Carter, C.S., 2000. Anterior cingulate and prefrontal cortex: who's in control? *Nat. Neurosci.* 3, 421–423.
- Cools, R., Clark, L., Owen, A.M., Robbins, T.W., 2002. Defining the neural mechanisms of probabilistic reversal learning using event-related functional magnetic resonance imaging. *J. Neurosci.* 22, 4563–4567.
- Damasio, A., 1999. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. Harcourt Brace and Co, New York.
- Damasio, A.R., Tranel, D., Damasio, H., 1990. Individuals with sociopathic behavior caused by frontal damage fail to respond automatically to social stimuli. *Behav. Brain Res.* 41, 81–94.
- Dolcos, F., LaBar, K., Cabeza, R., 2004. Interaction between the amygdala and the medial temporal lobe memory system predicts better memory for emotional events. *Neuron* 42, 855–863.
- Durston, S., Thomas, K.M., Worden, M.S., Yang, Y., Casey, B.J., 2002. The effect of preceding context on inhibition: an event-related fMRI study. *NeuroImage* 16, 449–453.
- Durston, S., Davidson, M.C., Thomas, K.M., Worden, M.S., Tottenham, N.T., Martinez, A., Watts, R., Ulug, A.M., Casey, B.J., 2003. Parametric manipulation of conflict and response competition using rapid mixed-trial event-related functional MRI. *NeuroImage* 20 (4), 2135–2141.
- Elliott, R., Dolan, R.J., Frith, C.D., 2000. Dissociable functions of the medial and lateral orbitofrontal cortex: evidence from neuroimaging studies. *Cereb. Cortex* 10, 308–317.
- Eslinger, P.J., Damasio, A.R., 1985. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* 35 (12), 1731–1741.
- Eslinger, P.J., Grattan, L.M., Damasio, H., Damasio, A.R., 1992. Developmental consequences of childhood frontal lobe damage. *Archives of Neurology* 49, 764–769.
- Everitt, B.J., Cardinal, R.N., Parkinson, J.A., Robbins, T.W., 2003. Appetitive behavior: impact of amygdala-dependent mechanisms of emotional learning. *Annual New York Academy of Sciences* 985, 233–250.
- Frith, C.D., Frith, U., 1999. Interacting minds: a biological basis. *Science* 286, 1692–1695.
- Garavan, H., Ross, T.J., Murphy, K., Roche, R.A., Stein, E.A., 2002. Dissociable executive functions in the dynamic control of behavior: inhibition, error detection, and correction. *NeuroImage* 17, 1820–1829.
- Grafman, J., Schwab, K., Warden, D., Pridgen, A., Brown, H.R., Salazar, A.M., 1996. Frontal lobe injuries, violence, and aggression: a report of the Vietnam Head Injury Study. *Neurology*, 465.
- Gray, N.S., MacCulloch, M.J., Smith, J., Morris, M., Snowden, R.J., 2003. Forensic psychology: violence viewed by psychopathic murderers. *Nature* 423, 497–498.
- Greene, J.D., Cohen, J.A., 2004. For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society B: Biological Sciences* 359, 1775–1785.
- Greene, J., Haidt, J., 2002. How and where does moral judgment work? *Trends in Cognitive Sciences* 6, 517–523.
- Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., Cohen, J.D., 2001. An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 1971–1972.
- Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D., 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44, 389–400.
- Greenwald, A.G., Banaji, M.R., 1995. Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review* 102, 4–27.
- Greenwald, A.G., Farnham, S.D., 2000. Using the Implicit Association Test to measure self-esteem and self-concept. *J. Pers. Soc. Psychol.* 79, 1022–1038.
- Greenwald, A.G., McGhee, D.E., Schwartz, K.L.K., 1998. Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology* 74, 1464–1480.
- Haidt, J., 2001. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review* 1084, 814–834.
- Hare, R.D., 1991. *The Hare Psychopathy Checklist-Revised*. Multi-Health Systems, Toronto, Ontario.
- Heekeren, H.R., Wartenburger, I., Schmidt, H., Schwintowski, H.P., Villringer, A., 2003. An fMRI study of simple ethical decision-making. *NeuroReport* 14, 1215–1219.
- Heekeren, H.R., Wartenburger, I., Schmidt, H., Prehn, K., Schwintowski, H.-P., Villringer, A., 2005. Influence of bodily harm on neural correlates of semantic and moral decision-making. *NeuroImage* 24 (3), 887–897.
- Helwig, C.C., Jasiobedzka, U., 2001. The relation between law and morality: children's reasoning about socially beneficial and unjust laws. *Child Dev.* 72, 1382–1393.
- Kagan, J., Lamb, S., 1987. *The Emergence of Morality in Young Children*. University of Chicago Press, Chicago.
- Kerns, J.G., Cohen, J.D., MacDonald III, A.W., Cho, R.Y., Stenger, V.A., Carter, C.S., 2004. What controls control? Anterior cingulate conflict monitoring predicts adjustments in control. *Science* 303, 1023–1026.
- Kohlberg, L., 1969. Stage and sequence: the cognitive-developmental approach to socialization. In: Goslin, D.A. (Ed.), *Handbook of Socialization Theory and Research*. Rand McNally, Chicago, pp. 347–480.
- Kohlberg, L., Kramer, R., 1969. Continuities and discontinuities in childhood and adult moral development. *Hum. Dev.* 12, 93–120.
- Kringelbach, M.L., Rolls, E.T., 2003. Neural correlates of rapid context-dependent reversal learning in a simple model of human social interaction. *NeuroImage* 20, 1371–1383.
- Lang, P.J., Greenwald, M.K., 1985. *The International Affective Picture System Slides and Technical Report*. Center for Research in Psychophysiology, University of Florida, Gainesville.
- Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1995. *International Affective Picture System IAPS*. National Institute of Mental Health Center for the Study of Emotion and Attention, Bethesda, MD.
- LeDoux, J., 1998. *The Emotional Brain*. Weidenfeld and Nicolson, New York.
- Leung, H.-C., Skudlarski, P., Gatenby, J.C., Peterson, B.S., Gore, J.C., 2000. An event-related functional MRI study of the Stroop color word interference task. *Cereb. Cortex* 10, 552–560.
- Liberzon, I., Taylor, S.F., Fig, L.M., Decker, L.R., Koeppe, R.A., Minoshima, S., 2000. Limbic activation and psychophysiological responses to aversive visual stimuli. Interaction with cognitive task. *Neuropsychopharmacology* 235, 508–516.
- Lykken, D.T., 1957. A study of anxiety in the sociopathic personality. *J. Abnorm. Soc. Psychol.* 55, 6–10.
- MacDonald III, A.W., Cohen, J.D., Stenger, V.A., Carter, C.S., 2000. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science* 288, 1835–1838.
- MacLeod, C.M., MacDonald, P.A., 2000. Interdimensional interference in the Stroop effect: uncovering the cognitive and neural anatomy of attention. *Trends Cogn. Sci.* 4 (10), 383–391.
- Menon, V., Adelman, N., White, C.D., Glover, G.H., Reiss, A.L., 2001. Error-related brain activation during a Go/NoGo response inhibition task. *Hum. Brain Mapp.* 123, 131–143.
- Milne, E., Grafman, J., 2001. Ventromedial prefrontal cortex lesions in humans eliminate implicit gender stereotyping. *J. Neurosci.* 21, RC150.
- Mink, J.W., 1996. The basal ganglia: focused selection and inhibition of competing motor programs. *Progr. Neurobiol.* 50, 81–425.
- Moll, J., Eslinger, P.J., Oliveira-Souza, R., 2001. Frontopolar and anterior temporal cortex activation in a moral judgment task: preliminary functional MRI results in normal subjects. *Arq. Neuropsiquiatr.* 59, 657–664.
- Moll, J., Oliveira-Souza, R., Eslinger, P.J., Bramati, I.E., Mourao-Miranda, J., Andreiuolo, P.A., Pessoa, L., 2002a. The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* 22, 2730–2736.
- Moll, J., Oliveira-Souza, R., Bramati, I.E., Grafman, J., 2002b. Functional networks in emotional moral and nonmoral social judgments. *NeuroImage* 16, 696–703.

- Moll, J., de Oliveira-Souza, R., Eslinger, P.J., 2003. Morals and the human brain: a working model. *NeuroReport* 14, 299–305.
- Nagahama, Y., Okada, T., Katsumi, Y., Hayashi, T., Yamauchi, H., Oyanagi, C., Konishi, J., Fukuyama, H., Shibasaki, H., 2001. Dissociable mechanisms of attentional control within the human prefrontal cortex. *Cereb. Cortex* 11, 85–92.
- Nichols, S., 2002. Norms with feeling: towards a psychological account of moral judgment. *Cognition* 84, 221–236.
- Nosek, B.A., Banaji, M.R., 2002. The Power of the Immediate Situation: Gender Differences in Implicit Math Attitudes. Society of Personality and Social Psychology, Savannah, GA.
- Nunez, J.M., Casey, B.J., Egner, T., Hare, T., Hirsch, J., 2005. Intentional false responding shares neural substrates with response conflict and cognitive control. *NeuroImage* 25, 267–277.
- Ochsner, K.N., Bunge, S.A., Gross, J.J., Gabrieli, J.D., 2002. Rethinking feelings: an fMRI study of the cognitive regulation of emotion. *J. Cogn. Neurosci.* 14, 1215–1229.
- O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., 2001. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102.
- O'Doherty, J., Critchley, H.D., Deichmann, R., Dolan, R.J., 2003. Dissociating valence of outcome from response switching in human orbital and ventral prefrontal cortices. *J. Neurosci.* 23, 7931–7939.
- Pardo, J.V., Pardo, P.J., Janer, K.W., Raichle, M.E., 1990. The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proc. Natl. Acad. Sci. U.S.A.* 87, 256–259.
- Pelphrey, K.A., Morris, J., McCarthy, G., 2004. Grasping the intentions of others: the perceived intention of an action influences activity in the superior temporal sulcus during social perception. *J. Cogn. Neurosci.* 16, 1706–1716. (Special issues on social neuroscience).
- Peterson, B.S., Kane, M.J., Alexander, G.M., Lacadie, C., Skudlarski, P., Leung, H.C., May, J., Gore, J.C., 2002. An event-related functional MRI study comparing interference effects in the Simon and Stroop tasks. *Brain Res. Cogn. Brain Res.* 13, 427–440.
- Phelps, E.A., O'Connor, K.J., Cunningham, W.A., Funayama, E.S., Gatenby, J.C., Gore, J.C., Banaji, M.R., 2000. Performance on indirect measures of race evaluation predicts amygdala activation. *J. Cogn. Neurosci.* 12, 729–738.
- Phelps, E.A., Cannistraci, C.J., Cunningham, W.A., 2003. Intact performance on an indirect measure of face bias following amygdala damage. *Neuropsychologia* 41, 203–208.
- Piaget, J., 1932. *The Moral Judgment of the Child*. Harcourt, New York.
- Purves, D., Augustine, G.J., Fitzpatrick, D., Katz, L.C., Lamantia, A.S., McNamara, J.O., Williams, S.M., 2001. *Neuroscience*. Sinauer Associates, Sunderland.
- Rolls, E.T., 2000. The orbitofrontal cortex and reward. *Cereb. Cortex* 10, 284–294.
- Rozin, P., Royzman, E.B., 2001. Negativity bias, negativity dominance, and contagion. *Pers. Soc. Psychol. Rev.* 5, 296–320.
- Ruff, C.C., Woodward, T.S., Laurens, K.R., Liddle, P.F., 2001. The role of the anterior cingulate cortex in conflict processing: evidence from reverse Stroop interference. *NeuroImage* 14, 1150–1158.
- Schall, J.D., Stuphorn, V., Brown, J.W., 2002. Monitoring and control of action by the frontal lobes. *Neuron* 36, 309–322.
- Talairach, J., Tournoux, P., 1988. *Co-planar Stereotaxic Atlas of the Human Brain*. Thieme, New York.