Brain (2002), **125,** 1696–1708

An fMRI study of intentional and unintentional (embarrassing) violations of social norms

S. Berthoz,[1,3] J. L. Armony,[1,2] R. J. R. Blair[1] and R. J. Dolan[2]

[1]*Institute of Cognitive Neuroscience, University College London,* [2]*Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK and* [3]*Department of Psychiatry, Institut Mutualiste Montsouris, Paris, France*

Correspondence to: Dr R. J. R. Blair, Institute of Cognitive Neuroscience, Alexandra House, 17 Queen Square, London WC1N 3BG, UK
E-mail: j.blair@ucl.ac.uk

Summary

The aim of this investigation was to identify neural systems supporting the processing of intentional and unintentional transgressions of social norms. Using event-related fMRI, we addressed this question by comparing neural responses to stories describing normal behaviour, embarrassing situations or violations of social norms. Processing transgressions of social norms involved systems previously reported to play a role in representing the mental states of others, namely medial prefrontal and temporal regions. In addition, the processing of transgressions of social norms involved systems previously found to respond to aversive emotional expressions (in particular angry expressions); namely lateral orbitofrontal cortex (Brodmann area 47) and medial prefrontal cortex. The observed responses were similar for both intentional and unintentional social norm violations, albeit more pronounced for the intentional norm violations. These data suggest that social behavioural problems in patients with frontal lobe lesions or fronto-temporal dementia may be a consequence of dysfunction within the systems identified in light of their possible role in processing whether particular social behaviours are, or are not, appropriate.

Keywords: social norm violations; embarrassment; Theory of Mind; orbitofrontal cortex; medial prefrontal cortex

Abbreviations: ToM = Theory of Mind

Introduction

Lateral orbitofrontal, ventromedial frontal and medial prefrontal cortex have been consistently linked to aspects of social cognition (e.g. Damasio, 1994; Baron-Cohen *et al*., 1999*b*; Frith and Frith, 1999; Blair and Cipolotti, 2000). For example, lesions to these regions have been associated with emotional and personality changes such as euphoria, irresponsibility, lack of affect and lack of concern for the present or future (Hecaen and Albert, 1978; Stuss and Benson, 1986). Furthermore, patients with aberrant behaviour following frontal lesions have been reported to show impairments in emotional expression recognition (Hornak *et al*., 1996; Blair and Cipolotti, 2000) and to perform poorly on self-reported measures of empathy (Grattan *et al*., 1994; Eslinger, 1998). In addition, frontal lobe damage has been linked to impairments in social behaviour. Patients with frontal lobe lesions have been described as presenting diminished social awareness and a lack of concern for social rules (e.g. Lishman, 1968; Blumer and Benson, 1975; Hecaen and Albert, 1978; Stuss *et al*., 1992; Damasio, 1994). Frequently, increased levels of aggression and aberrant behaviour are reported, and these are found both when lesions are acquired early in life as well as in adulthood (e.g. Burgess and Wood, 1990; Price *et al*., 1990; Pennington and Bennetto, 1993; Damasio, 1994; Volavka, 1995; Grafman *et al*., 1996; Anderson *et al*., 1999).

It has been suggested that some of the social difficulties faced by patients with ventromedial frontal cortex damage reflect an impairment in the system responsible for the representation of the mental states of others, i.e. Theory of Mind (ToM) (Baron-Cohen, 1995). In addition, a series of neuroimaging studies in healthy individuals that, using a variety of different paradigms, have consistently implicated regions of medial prefrontal cortex in the representation of the mental states of others (Fletcher *et al*., 1995; Goel *et al*., 1995; Baron-Cohen *et al*., 1999*b*; Brunet *et al*., 2000; Castelli *et al*., 2000; Gallagher *et al*., 2000; Vogeley *et al*., 2001), as well as other regions, including left temporo-parietal regions and the temporal poles (for a review see Frith and Frith, 1999). Moreover, there are now several reports of patients with lesions to medial prefrontal cortex who show ToM impairment (Happe *et al*., 2001; Lough *et al*., 2001; Stuss *et al*., 2001; Rowe *et al*., 2001). Importantly, neuropsychological work has demonstrated that ToM is doubly dissociable from other executive processes that rely on prefrontal cortex, such as working memory and impulse control (Blair and Cipolotti, 2000; Fine *et al*., 2001; Lough *et al*., 2001).

It should be noted, however, that not all patients exhibiting social behavioural difficulties following lesions of frontal cortex present with impairments in ToM (e.g. Blair and Cipolotti, 2000; cf. Saver and Damasio, 1991). Studies of these patients' social behavioural difficulties have suggested at least two affect-based accounts of their impairment (Damasio, 1994; Blair and Cipolotti, 2000). One influential account explains the patient's social difficulties in terms of damage to the somatic marker system (Damasio, 1994; Bechara *et al*., 2000*a*, *b*). In this account, somatic markers provide signals of the inappropriateness of particular behaviours allowing their rejection. The absence of these signals prevents inappropriate courses of action from being rejected and leads to behavioural disturbance.

An alternative account of the social behavioural difficulties following frontal lesions stresses the role of social cues in modulating social behaviour (Blair and Cipolotti, 2000; Blair, 2001), by reference to the role of orbitofrontal cortex in response reversal as a consequence of changes in reinforcement contingencies (Dias *et al*., 1996; Rolls, 2000). Angry expressions are known to curtail the behaviour of others in situations where social rules or expectations have been violated (Averill, 1982). The suggestion is that activation of a distinct system by angry expressions or expectations of another's anger results in the modulation of current behavioural response in the individual engaging in inappropriate behaviour. This hypothesis has drawn support from findings that ventrolateral orbitofrontal cortex [Brodmann area (BA) 47] is activated by negative emotional expressions; in particular anger, but also fear and disgust (Sprengelmeyer *et al*., 1998; Blair *et al*., 1999; Kesler-West *et al*., 2001). Moreover, patients with social behavioural difficulties following lesions of frontal cortex, particularly the orbitofrontal region, are also impaired in the recognition of facial expressions, particularly anger (Hornak *et al*., 1996; Blair and Cipolotti, 2000). Such patients have also been found to show impairment in appropriately attributing anger and embarrassment, but not other emotions, to story protagonists, and, in addition, to have a deficit in identifying violations of social norms (Stone *et al*., 1998; Blair and Cipolotti, 2000).

Importantly, this form of social response reversal appears to be mediated by a system that can be dissociated both neuropsychologically and pharmacologically from the neural circuitry implicated in response reversal as a function of changes in reinforcement schedules. Thus, neurological patients have been found who present with severe impairment in social response reversal but no impairment on tasks of response reversal as a function of contingency change (Blair and Cipolotti, 2000). In contrast, adult psychopathic individuals show pronounced impairment in response reversal to contingency change, but no impairment on tasks purported to index social response reversal (Blair and Cipolotti, 2000; Mitchell *et al*., 2002). Additionally, GABAergic compounds such as alcohol and diazepam interfere with social response reversal (Borrill *et al*., 1987; Blair and Curran, 1999), but not response reversal to contingency change (Coull *et al*., 1995). In contrast, serotonergic manipulations modulate response reversal to contingency change (Rogers *et al*., 1999), but not social response reversal (Harmer *et al*., 2001*a*).

The processes underlying the ability to identify social norm violations are of interest in that an impairment is likely to be associated with severe social behavioural difficulties (Dewey, 1991; Stone *et al*., 1998; Blair and Cipolotti, 2000). In order to determine whether another individual has violated social norms, it is necessary to calculate whether the action could be construed as a violation and, crucially, whether the action was intended. Actions that could be construed as violations, but which are unintentional, are generally considered embarrassing rather than social violations (Garland and Brown, 1972; Semin and Manstead, 1982).

Although accounts of the cognitive processes underlying the ability to identify norm violations have been proposed, the brain systems involved in mediating these processes remain unknown. This is despite the fact that this information is likely to be crucial for understanding the social behavioural difficulties of patients with frontal brain lesions. Based on the account presented above, we can hypothesize that brain regions involved in the representation of the mental states of others will be activated during violation situations. These include medial prefrontal cortex, temporo-parietal regions and temporal pole (for a review see Frith and Frith, 1999). In addition, we suggest that processing social violations will also involve systems associated with the representation of aversive emotional reactions in others, particularly others' anger. These areas include lateral orbitofrontal (BA 47) and medial prefrontal cortices (Sprengelmeyer *et al*., 1998; Blair *et al*., 1999; Harmer *et al*., 2001*b*; Kesler-West *et al*., 2001).

Material and methods

Participants

Twelve right-handed males, with a mean age of 26 ± 5 years (range 19–37 years), free of past and present neurological disorder, participated in this study. Participants were required to have English as their first language. This study was approved by the Institute of Neurology Ethics Committee. Informed written consent was obtained from all subjects prior to scanning.

Task

Four types of verbal material were presented. These were stories with three types of endings, and sentences of 'unrelated words'. The endings were either: (i) description of a normal situation (normal: normative social behaviour); (ii) description of an embarrassing situation for the story protagonist (embarrassment: unintentional transgression); or (iii) description of a situation where the story protagonist's behaviour is a violation of social norms (violation: intentional transgression). The difference between the embarrassing and violation issues was intention, the former representing unintentional and the latter intentional social violations.

The four types of verbal material were presented twice: once with a personal reference (i.e. the story protagonist is 'you') and once with an impersonal reference (i.e. the story protagonist is a character). In total, there were 120 stories (20 normal-personal, 20 normal-impersonal; 20 embarrassing-personal, 20 embarrassing-impersonal; 20 violation-personal, 20 violation-impersonal) and 40 unrelated words (20 personal, 20 impersonal).

Examples

Personal stories

Beginning. 'You are invited for a Japanese dinner at your friend's house'.
Endings. (i) Normal: 'You have a bite of the first course and like it, and congratulate your friend for her good cooking'. (ii) Embarrassing: 'You have a bite of the first course, you choke and spit out the food while you are coughing'. (iii) Violation: 'You have a bite of the first course, but do not like it and spit the food back into your plate'.

Impersonal stories

Beginning. 'Joanna is invited for a Japanese dinner at her friend's house'.
Endings. (i) Normal: 'She has a bite of the first course and likes it, and congratulates her friend for her good cooking'. (ii) Embarrassing: 'She has a bite of the first course, chokes and spits out the food while she is coughing'. (iii) Violation: 'She has a bite of the first course, but does not like it and spits the food back into her plate'.

Personal unrelated words

Beginning. 'You in changed oil little a when had two the since'.
Ending. 'A in already then uncle day you of performance his come end and put'.

Impersonal unrelated words

Beginning. 'Paul in changed oil little a when had two the since'.
Ending. 'A in already then uncle day Paul of performance his come end and put'.

The stories used in this task have been validated in two previous experiments (S. Berthoz, T.Farquhar, R.J.R.Blair, unpublished results). Results of these experiments showed that, first, subjective ratings of embarrassment and inappropriateness of, respectively, the embarrassing and social violation stories were significantly greater than the ones of the normal stories. Furthermore, the ratings of embarrassment were significantly greater for the embarrassing than for the social violation stories, and the ratings of inappropriateness of behaviour were significantly greater for the social violation than for the embarrassing stories. Secondly, autonomic activity, as indexed by skin conductance response (SCR), was seen to be significantly greater to intentional and unintentional violations in comparison with neutral stories.

Procedure

All stimuli were displayed on a monitor and presented to the participant via a 45° angled mirror positioned above the head coil. This mirror was adjusted to be within the participant's field of vision without having to tilt the head. A test image was presented on the screen prior to scanning to ensure that the image was in focus and the participant could comfortably read the text.

The beginning of the story was presented on the screen for 8 s. This text was then replaced by the end of the story, which was presented for 10 s. The stories were separated by a 1-s grey screen. The experiment comprised two sessions of 80 stimuli (60 stories and 20 unrelated words), presented in pseudo-random order. Participants were told they would see some stories on the computer screen, and that these stories would be repeated with various types of endings. In addition, they were told the stories would either describe what was happening to a character or what was happening to 'you'. Participants were instructed to read the text silently, and to click the response key when they finished reading the second part of the story. They were instructed to try to imagine what they/the story protagonist would feel if they were in the situation described. The participants were also told they would see some sentences that were artificially created by mixing unrelated words together, and that these sentences had no meaning. Participants were advised to simply read them, without trying to extract any meaning.

Following the scanning session, the participants were asked to rate, for all the different stories: (i) how embarrassing they thought the situation to be; (ii) how inappropriate they thought the behaviour to be; (iii) how funny they thought the story to be, using three separate seven-point Likert scales ranging from 1 (i.e. 'indifferent', 'understandable' and 'not funny at all', respectively) to 7 (i.e. 'extremely embarrassed', 'completely inappropriate' and 'extremely funny', respectively). Half of the participants rated the personal stories, and the other half rated the impersonal stories.

Data acquisition

Data were acquired on a 2T Siemens VISION whole-body MRI system equipped with a head volume coil. Functional (T_2*-weighted) echoplanar image volumes were acquired using blood oxygenation level dependency (BOLD) contrast. A total of 1020 images (510 images per run) were taken for each participant, each comprising a full brain volume of 40 contiguous axial slices (1.8-mm thickness). Volumes were acquired continuously with an effective repetition time (TR) of 3.04 s. A total of 1008 images per subject were analysed (five dummy volumes at the beginning of each run and one dummy volume at the end of each run were discarded). A T_1-weighted anatomical MRI was also acquired for each subject.

Data analysis

Processing and analysis of the functional (T_2*-weighted) images was performed using statistical parametric mapping (SPM; version SPM99 was used) (Friston *et al.*, 1994; Worsley and Friston, 1995; see also http://www.fil.ion.ucl.ac.uk/spm/). Images were realigned to the first volume of each session to correct for interscan movement and spatially normalized to standard Talairach space using a template provided by the Montreal Neurological Institute (Evans *et al.*, 1994). Finally, images were spatially smoothed using an 8 mm (full width half maximal) Gaussian kernel. The evoked responses for the eight experimental conditions were modelled using a boxcar of 10-s duration convolved with a synthetic haemodynamic response function (hrf). Key presses were modelled using a delta function convolved with the synthetic hrf. The six movement parameters obtained from the realignment procedure were also included in the model as confounds to minimize the chances of detecting false activations due to movement artefacts.

Data were analysed using a two-level mixed effects model (random effect analysis): linear contrasts of the parameter estimates for the effects of interest were obtained for each subject (combining the two sessions) and then entered into a between-subject one-sample *t*-test. The resulting *t* statistic at every voxel constitutes a statistical parametric map. Results are presented with a threshold at $P < 0.0001$ (uncorrected). To minimize the risk of type I errors, a correction for spatial extent ($P < 0.05$) was applied (Friston *et al.*, 1994); functionally, this meant that all clusters were >15 voxels in size.

To determine the common areas activated by the embarrassment and violation conditions, the SPM T-map ensuing from the contrast of embarrassing minus normal stories in the one-sample *t*-test was thresholded at 3.09 ($P < 0.001$) to create a binarized mask of this contrast. This mask was then applied to the contrast of violation minus normal stories.

Results

Subjective rating

Repeated-measures ANOVA (analysis of variance) with two within-subjects factors (Type of Story: normal, embarrassing, violation; Ratings: embarrassment, inappropriateness, humour), and one between-subjects factor (Perspective: personal, impersonal) were performed. No difference between the personal and impersonal ratings was found. Significant main effects of Type of Story [$F(2,20) = 187.98$; $P < 0.0001$], and of Ratings [$F(2,20) = 18.91$; $P < 0.0001$], and a significant Type of Story × Ratings interaction [$F(4,40) = 28.04$; $P < 0.0001$] were shown. Planned comparisons (two-tailed *t*-test) revealed a significant difference between the mean ratings of inappropriateness of behaviour for normal and social violations stories [$t(11) = 15.21$; $P < 0.0001$], and a significant difference between the mean ratings of embarrassment for normal and embarrassing stories [$t(11) = 19.47$; $P < 0.0001$]. Moreover, mean ratings of the inappropriateness of behaviour were significantly greater for the social violations than for the embarrassing stories [$t(11) = 8.26$; $P < 0.0001$], and mean embarrassment ratings were significantly greater for the embarrassing than for the social violations stories [$t(11) = 3.25$; $P < 0.01$]. Finally, social violation and embarrassing stories were rated as funnier than the normal ones [$t(11) = 5.65$; $P < 0.0001$; and $t(11) = 6.59$; $P < 0.0001$, respectively], but social violation and embarrassing stories did not differ significantly and were rated as similarly funny.

fMRI data

The main objective of the present study was to investigate two potentially dissociable systems for social cognition, one involved in the processing of embarrassment (unintentional) and the other in the processing of intentional violation of social norms. We addressed this by comparing neural responses to stories describing either a normal behaviour, an embarrassing situation, or violation of social norms. In the present analyses, we used the normal stories as a reference condition, and collapsed the personal and impersonal stories.

Violation of social norms versus normal behaviour

Comparison of brain activity associated with the violation of social norm stories and activity associated with normal stories showed foci of significant increased activation in medial and superior (bilaterally) prefrontal cortex, left middle and inferior prefrontal cortex, left orbitofrontal cortex, anterior temporal pole bilaterally, left temporo-parietal junction, occipital cortex with foci in cuneus and posterior fusiform gyrus, and the brainstem (exact coordinates are given in Table 1 and Fig. 1).

Embarrassment versus normal behaviour

Comparison of brain activity associated with the embarrassing stories and activity associated with normal stories showed revealed foci of significant increased activation in the right medial and superior prefrontal cortex, left middle and inferior prefrontal cortex, left orbitofrontal cortex, anterior and middle temporal pole bilaterally, left temporo-parietal junction and occipital cortex with foci in cuneus and fusiform gyrus (exact coordinates are given in Table 1 and Fig. 2).

Table 1 *Brain activity related to intentional and unintentional violation of social norms relative to normal social behaviours*

Region (BA)	*Z*-score	Coordinates			Cluster size
		x	*y*	*z*	
Intentional violation of social norms minus normal stories					
L medial frontal (10)	4.95	–12	58	8	192
L medial frontal (9)	4.79	–8	52	18	
R medial frontal (9)	4.04	10	54	24	
L precentral (4)	4.46	–10	–18	66	41
L middle frontal (6)	4.07	–44	6	44	50
R superior frontal (8)	5.10	6	36	54	277
L superior frontal (8)	4.56	–6	32	54	
L superior frontal (6)	3.74	–6	14	60	38
L dorsolateral prefrontal cortex (45)	4.31	–52	32	2	72
L lateral orbitofrontal cortex (47)	3.47	–36	28	–22	43
L middle temporal (21/37)	4.02	–50	–44	6	16
R middle/superior temporal (21/38)	4.25	56	8	–28	107
L middle/superior temporal (21/38)	5.19	–52	8	–38	245
L cuneus (17)	4.79	–8	–98	8	686
L fusiform gyrus (19/18)	4.32	–20	–82	–10	
R cuneus (18)	4.23	18	–98	10	68
L cerebellum	3.76	–34	–74	–16	31
Ponto-mesencephalic junction	3.78	–6	–28	–20	23
Unintentional violation (embarrassing) violation of social norms minus normal stories					
R medial frontal (9)	4.30	4	54	36	64
R superior frontal (8)	3.77	2	36	52	20
L middle frontal (6/9)	4.19	–48	12	46	203
L dorsolateral prefrontal cortex (44)	4.25	–56	18	16	
L dorsolateral prefrontal cortex (45)	4.41	–54	24	8	184
L lateral orbitofrontal cortex (47)	4.21	–42	26	–14	
L middle temporal (21/37)	3.95	–50	–38	–2	111
R middle temporal (21)	3.95	52	–16	–16	78
L middle temporal (21)	5.62	–54	2	–28	244
L superior temporal (38)	3.87	–50	12	–24	
R superior temporal (38)	4.74	48	8	–28	211
L lingual (18)	4.35	–4	–80	–4	444
L fusiform (18)	3.93	–30	–80	–12	
L cuneus (17)	4.69	–8	–102	6	201
R cuneus (17/18)	4.24	20	–96	10	142

L = left; R = right.

Violation of social norms versus embarrassing stories

Significant differential activation was seen in left medial and superior prefrontal cortex, anterior cingulate gyrus, precentral/postcentral sulcus, right temporal pole, left inferior parietal cortex, and left superior occipital gyrus and precuneus (exact coordinates are given in Table 2 and Fig. 3).

Embarrassing versus violation of social norms stories

The only significant focus of differential increased activity was seen in right temporal pole (BA 21) (exact coordinates are given in Table 2 and Fig. 4).

Commonalties of violation of social norms and embarrassing stories

Both violation of social norms and embarrassing stories activated left medial, middle and inferior prefrontal gyrus, superior frontal gyrus bilaterally, left orbitofrontal cortex, anterior and middle temporal pole bilaterally, left temporo-parietal junction and occipital cortex with foci in the cuneus, lingual and posterior fusiform gyri (exact coordinates are given in Table 3).

Discussion

Knowledge of the neural systems involved in processing violations of social norms is likely to be crucial for understanding the social behavioural difficulties of some

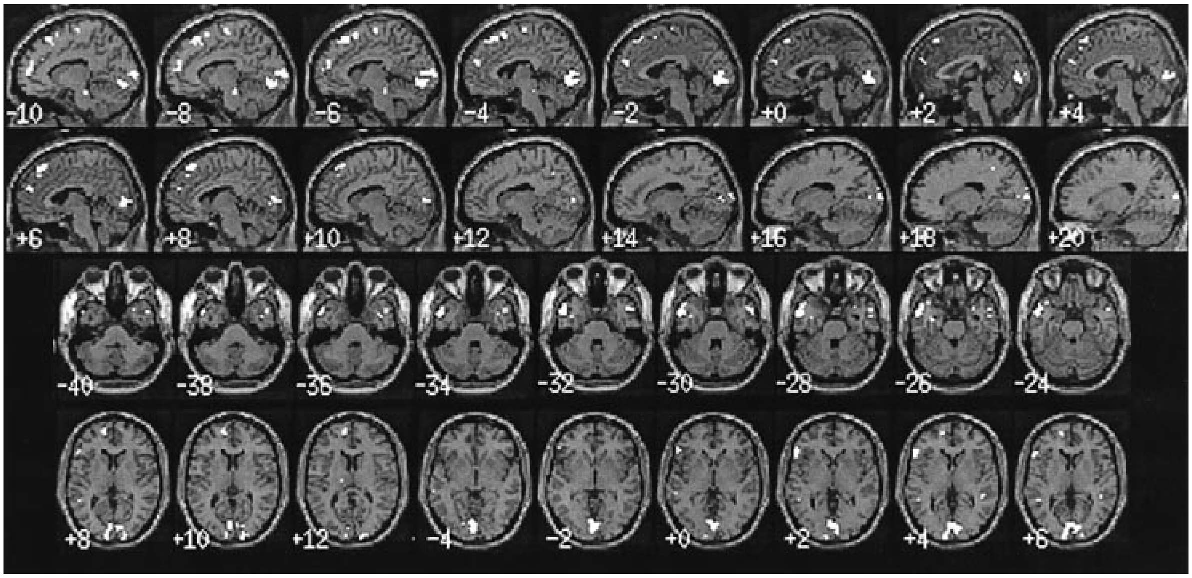

Fig. 1 Intentional violation of social norms minus normal social behaviour. Comparison of the intentional violation of social norms and normal behaviour conditions, with a height threshold set to $P < 0.001$ (uncorrected) and an extent threshold set to $P < 0.05$. Activations data are superimposed onto the canonical template of SPM99 in the sagittal (two top rows) and axial (two bottom rows) planes. x and z coordinates (for the sagittal and axial planes respectively) in the Talairach and Tournoux space are given (cf. Table 1).

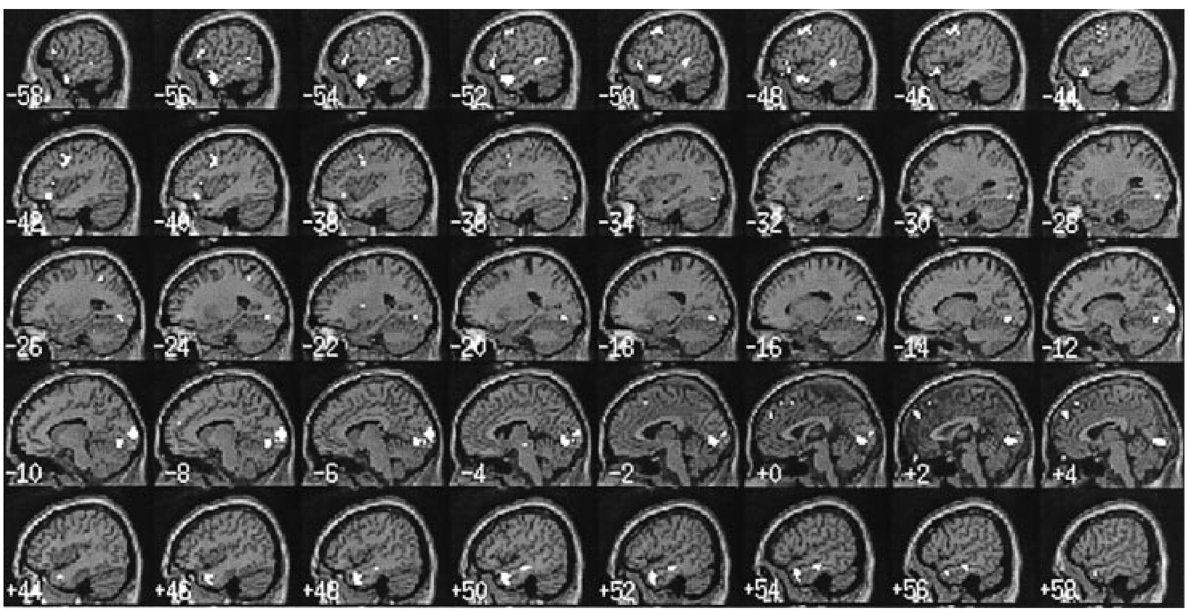

Fig. 2 Unintentional violation of social norms minus normal social behaviour. Comparison of the unintentional violation of social norms (embarrassment) and normal behaviour conditions, with a height threshold set to $P < 0.001$ (uncorrected) and an extent threshold set to $P < 0.05$. Activations data are superimposed onto the canonical template of SPM99 in the sagittal plane. x coordinates in the Talairach and Tournoux space are given (cf. Table 1).

patients with frontal brain lesions. Previous studies have examined the neural correlates of imagined aggressive behaviour (Pietrini *et al*., 2000), empathy and forgiveness (Farrow *et al*., 2001), and ToM (for a review see Frith and Frith, 1999). However, as far as we are aware, this is the first study to investigate the neural systems involved in the response to social norm violations, both intentional and unintentional (embarrassment).

The results of the current study are consistent with our primary hypotheses and showed that processing of violations of social norms involved systems previously reported to play a role in the representation of the mental states of others; namely, medial prefrontal and temporal areas (Frith and Frith, 1999). Furthermore, the neural response to violations of social norms involved systems previously found to respond to aversive emotional reactions in others (in particular others'

Table 2 *Comparisons between regional brain activity related to intentional and unintentional violation of social norms*

Region (BA)	Z-score	Coordinates			Cluster size
		x	*y*	*z*	
Intentional violation of social norms minus unintentional violation (embarrassing) stories					
L superior frontal (8)	4.15	–22	34	48	211
L superior frontal (6)	4.12	–24	24	60	
L medial frontal (10)	4.27	–10	58	8	66
L medial/ACing. (10/32)	3.84	–18	48	4	18
Cingulate (24)	3.88	0	0	40	43
Slcs GpoC/GprC (4)	4.21	–32	–22	42	39
L inferior parietal lobe (40)	4.93	–44	–32	52	113
R superior temporal (38)	4.30	36	–6	–32	16
L precuneus (7)	3.60	12	–50	52	19
L occipital superior (19)	4.14	–40	–72	32	85
Embarrassing minus violation of social norms stories					
R inferior temporal gyrus (21)	3.68	58	–18	–18	28

L = left; R = right; ACing. = anterior cingulate; Slcs GpoC/GprC = precentral/postcentral sulcus.

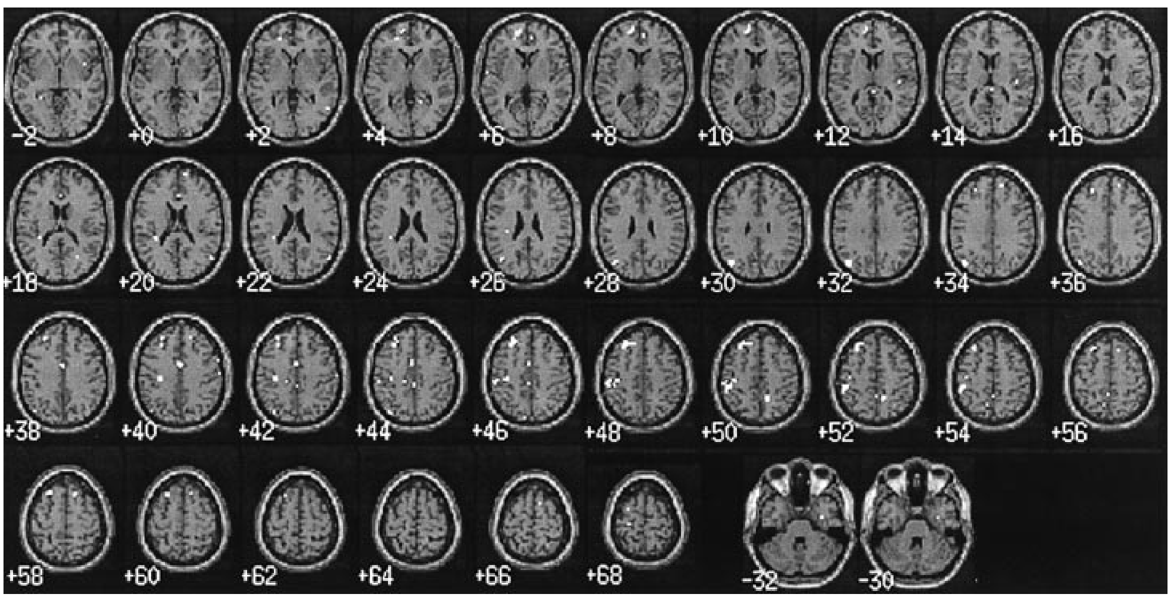

Fig. 3 Intentional minus unintentional violation of social norms. Comparison of the intentional and unintentional violation of social norms conditions, with a height threshold set to $P < 0.001$ (uncorrected) and an extent threshold set to $P < 0.05$. Activations data are superimposed onto the canonical template of SPM99 in the axial plane. *z* coordinates in the Talairach and Tournoux space are given (cf. Table 2).

anger): the lateral orbitofrontal cortex (BA 47) and the medial prefrontal cortex (Sprengelmeyer *et al.*, 1998; Blair *et al.*, 1999; Kesler-West *et al.*, 2001). The possible role(s) of the neural structures identified in this study is discussed in more detail below.

Medial prefrontal cortex

We observed activation of several regions within medial prefrontal cortex (BA 6, 8 and 9) when comparing brain activity related to violations of social norms and embarrassment conditions with that associated with normative social behaviours. This pattern of medial prefrontal response parallels the results of earlier functional imaging studies investigating the neural correlates of ToM. To illustrate, Goel *et al.* (1995) found BA 9 activation to be associated with reasoning about other people's thoughts regarding a novel object. Fletcher *et al.* (1995) found BA 8 activation to be associated with inferences about a character's intentions during a story comprehension task. Gallagher *et al.* (2000),

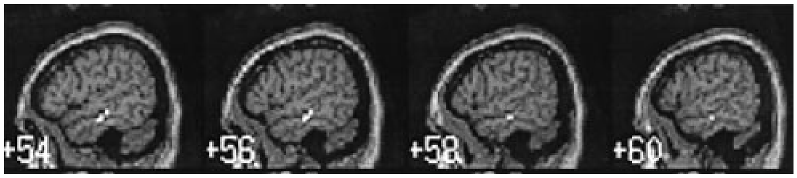

Fig. 4 Unintentional minus intentional violation of social norms. Comparison of the unintentional and intentional violation of social norms conditions, with a height threshold set to $P < 0.001$ (uncorrected) and an extent threshold set to $P < 0.05$. Activations data are superimposed onto the canonical template of SPM99 in the sagittal plane. x coordinates in the Talairach and Tournoux space are given (cf. Table 2).

Table 3 *Commonalties of regional brain activity related to intentional and unintentional violation of social norms*

Region (BA)	*Z*-score	Coordinates		
		x	*y*	*z*
L medial frontal (9)	4.79	–8	52	18
L middle frontal (6)	4.07	–44	6	44
L middle frontal (9)	3.56	–38	4	38
R superior frontal (8)	5.08	6	38	52
L superior frontal (8)	3.67	–4	40	52
L superior frontal (6)	3.26	–4	18	62
L dorsolateral PFC (45)	4.31	–52	32	2
L dorsolateral PFC (46)	3.15	–56	24	16
L lateral OFC (47)	3.53	–46	24	–16
L lateral OFC (47)	3.21	–38	28	–12
L middle temporal (22)	4.02	–50	–44	6
L middle temporal (21)	3.12	–56	–34	0
R middle temporal (21)	3.33	54	0	–28
R superior temporal (38)	4.25	56	8	–28
L superior temporal (38)	5.19	–52	8	–38
R cuneus (18)	4.23	18	–98	10
L cuneus (18)	4.79	–8	–98	8
L lingual G (18)	4.32	–20	–82	–10
R cuneus (17)	4.18	8	–90	6
L fusiform G (18)	4.10	–36	–84	–16

L = left; R = right; PFC = prefrontal cortex; OFC = orbitofrontal cortex.

using the same story comprehension task and a comparable non-verbal comprehension task involving static single-frame cartoons, found a convergence between activations that included BA 6 and 8/9 in response to verbal and visual stimuli that prompt mental state attributions. Brunet *et al.* (2000) found activation of BA 6 and 8/9 in participants viewing comic scripts that required the participant to attribute intentions to story characters. Castelli *et al.* (2000) found activation of BA 9 in healthy individuals when watching silent animations involving geometric shape characters that acted in a way to encourage the attribution of intentions.

The activation of these regions of medial prefrontal cortex in conjunction with the regions of temporal cortex strongly suggests that neural systems involved in the representation of the mental states of others are also involved in processing violations of social norms. This suggestion is also in line with neuropsychological data indicating that patients with impairments in the representation of the mental states of others, in particular individuals with autism, show difficulty in tasks that require the identification of norm violations or social *faux pas* (Dewey, 1991; Baron-Cohen *et al.*, 1999*a*). We propose that when one is confronted with the stimulus of another individual engaging in an act that will cause social disapproval, ToM is engaged. This engagement allows an observer to determine whether the individual is intentionally engaging in the act (and thus in need of social disapproval) or unintentionally engaging in the act (and thus expected to display signals of embarrassment).

In this context, it is interesting to note that many of these medial frontal cortical regions are activated to a greater extent by violations of social norms than by embarrassing situations (see Table 2). It might be argued that this difference is due to an arousal confound. There are recent reports relating medial prefrontal cortical activity to the generation of autonomic activity (Critchley *et al.*, 2000, 2001). The greater medial frontal cortical activity to the social norm violations rather than the embarrassing situations might reflect greater autonomic responses to the intentional norm violations. However, the regions identified in the studies investigating the generation of autonomic activity include anterior cingulate cortex and not the more anterior regions identified in the present study. Moreover, in our behavioural work (S. Berthoz, T.Farquhar, R.J.R.Blair, unpublished results), we found that while participants did show autonomic responses to the intentional social norm violations, they also showed autonomic responses of equivalent magnitude to the embarrassment/unintentional condition. Thus, the greater activity of medial prefrontal regions to the social norm violations is unlikely to be simply due to a difference in the level of arousal between the intentional and unintentional conditions.

It is possible that the greater medial frontal cortical activity to the violations of social norms may reflect an increased computational load for intentional violations relative to the embarrassment/unintentional conditions. In the embarrassment condition, it is only necessary to represent that the protagonist did not do the action intentionally. This would still require more activity than the normative conditions, where no intention needs to be calculated. However, in the violation of social norm stories, not only do the participants have to represent that the action is intentional, they are also likely to attempt to represent what that intention may be. The greater medial prefrontal activity in the intentional norm violations condition relative to the embarrassment/unintentional condition may possibly reflect an attempt to determine the story protagonist's actual intention.

In the current study, and in previous studies of ToM, activity in medial prefrontal cortex lies at the border of anterior cingulate cortex and medial frontal cortex in the paracingulate sulcus (Goel *et al*., 1995; Castelli *et al*., 2000; Gallagher *et al*., 2000). Two recent reviews of functional imaging studies that have activated anterior cingulate cortex have concluded that there is a degree of functional specialization in this region. Tasks that are predominantly cognitive engage the posterior part, while the anterior part shows greater activation when emotions are involved (Paus *et al*., 1998; Bush *et al*., 2000). In neuroimaging studies of ToM, it is not obvious why an area of cortex linked to emotional processing should be activated by tasks requiring the representation of mental states. However, it may be because the representation that another, or the self, has engaged in an intentional norm violation or a social, embarrassing *faux pas* has immediate emotional consequences.

Temporo-parietal region

Differential activation of the left temporo-parietal region was observed when comparing responses to violations of social norms and embarrassment situations relative to normative social behaviours. This region has frequently been observed in studies investigating ToM (Brunet *et al*., 2000; Castelli *et al*., 2000; Gallagher *et al*., 2000). This area has also been found to be highly sensitive to rotational movement, as well as visual and biological motion (Perrett *et al*., 1985; Bonda *et al*., 1996; Oram and Perrett, 1996; Puce *et al*., 1998; Grezes *et al*., 1999). In addition, human faces and animals, including animals without faces, activate this region (Chao *et al*., 1999). It has been suggested that the primary role of this region, as part of the ToM system, is the detection of animacy cues which, when detected, orchestrate the rest of the system to determine the intention of the detected animate being (Cipolotti *et al*., 1999; Blair *et al*., 2002).

Basal temporal cortex

In addition to activation of left temporo-parietal region, we observed bilateral activation in basal temporal regions. Again, this activation was significant for both violations of social norms and embarrassment conditions, relative to normative social behaviours. Similar patterns of activation have been frequently observed in studies investigating ToM (Goel *et al*., 1995; Brunet *et al*., 2000; Castelli *et al*., 2000; Gallagher *et al*., 2000). Moreover, these regions have also been linked to biological movement perception (Bonda *et al*., 1996; Grezes *et al*., 1999).

The finding of activation of the left temporo-parietal region and bilateral activation in the temporal poles to both violations of social norms and embarrassment conditions supports our suggestion that the neural systems involved in ToM are involved in processing social norm transgressions. Moreover, it is interesting to note that not only the medial prefrontal cortical regions, but also the left parietal region and right temporal pole, showed greater responses to the intentional violations of social norms relative to the embarrassment/unintentional conditions. This provides evidence in favour of the suggestion that this additional activity may reflect the individual's attempt to determine the story protagonist's actual intention for engaging in the social norm violation. For the embarrassment conditions, it is only necessary to represent that the individual did not have an intention. Interestingly, at least two predictions can be generated from these results. First, it can be predicted that patients with lesions to these temporal regions should show impairment on ToM tasks. Secondly, and more importantly, it can be predicted that patients with lesions to these regions will show impairment in responding appropriately to social norms. This is of particular relevance with respect to patients with frontotemporal dementia. Frontotemporal dementia is the term now preferred to describe the group of non-Alzheimer neuro-degenerative conditions affecting the frontal/temporal lobes (Lund and Manchester Groups, 1994; Neary *et al*., 1998). Patients with frontotemporal dementia frequently exhibit inappropriate social behaviour (Gregory and Hodges, 1996; Miller *et al*., 1997). It is plausible that for some patients with frontotemporal dementia, the disturbance in social functioning is due to disruption of temporal systems that are crucial for the maintenance of appropriate social cognition.

Orbitofrontal cortex

Clinical observations in humans and experimental reports in primates have consistently indicated that prefrontal cortex, in particular orbitofrontal cortex, is engaged in the regulation of social and aggressive behaviour (Damasio, 1994; Grafman *et al*., 1996; Blair and Cipolotti, 2000; Davidson *et al*., 2000; Rolls, 2000; Pietrini *et al*., 2000). Moreover, recent studies have shown impoverished prefrontal cortex functioning or prefrontal structural abnormalities in violent, antisocial adults (Volkow and Tancredi, 1987; Raine *et al*., 2000).

As expected, results of the present study showed substantial activation of left orbitofrontal cortex (BA 10 and 47) associated with both intentional and unintentional violations of social norms stories. It is also worth noting that the region of BA 47 activated in this study is almost identical to the region previously found to be activated by angry expressions (Sprengelmeyer *et al*., 1998; Kesler-West *et al*., 2001) [although Blair and colleagues observed the same region of orbitofrontal cortex response to angry faces, but in the right hemisphere (Blair *et al*., 1999)]. This region is also activated under conditions when an individual is induced to feel angry

(Dougherty *et al*., 1999). Thus, this region of BA 47 not only responds to the angry expressions of others, but also responds to stimuli detailing actions that are likely to cause others to become angry. Activation of this region might result in modulation of current behavioural responding to prevent the individual from engaging in inappropriate behaviour. When representing a social norm violation or embarrassment situation, the individual forms an expectation of other's anger/social disapproval that reverses this response option in favour of another (Blair and Cipolotti, 2000; Blair, 2001). The suggestion is that damage to this system will lead to socially inappropriate behaviour and, possibly, reactive aggression as the patient considering an inappropriate action will not reverse their decision on the basis of the expected social disapproval.

The amygdala

The amygdala was not found to be involved in the processing of intentional and unintentional social transgressions. We believe that this is because the amygdala is not engaged in the form of social decision making indexed by the task used in the current study (Blair and Cipolotti, 2000; Blair, 2001). Indeed, information about others' social disapproval/anger is thought to be received from sensory association cortices rather than the amygdala (Blair, 2001). In the present experiment, the regions of orbitofrontal cortex that were activated by the social norm violations and embarrassment stories are known to receive distinct input from sensory association cortices (e.g. Carmichael and Price, 1995). Furthermore, to date, no functional imaging studies have identified amygdala activation as part of the neural response to angry expressions (Sprengelmeyer *et al*., 1998; Blair *et al*., 1999; Kesler-West *et al*., 2001). Moreover, psychopathic individuals do not exhibit impairment in social decision making (Blair and Cipolotti, 2000; Blair *et al*., 2001), but have pronounced deficit on functions that rely on the amygdala such as: aversive conditioning (Hare and Quinn, 1971), augmentation of the startle reflex (Levenston *et al*., 2000) and recognition of fearful expressions (Blair *et al*., 2001; Stevens *et al*., 2001).

In another respect, there have been some suggestions that the amygdala may play a role in either the development, or functioning, of ToM (Baron-Cohen *et al*., 1999*b*, 2000; Fine *et al*., 2001). In the current study, we observed the expected activation of those brain regions linked to ToM, yet we did not observe activation of the amygdala. However, it is worth mentioning that, among all the functional imaging studies of ToM (Fletcher *et al*., 1995; Goel *et al*., 1995; Brunet *et al*., 2000; Castelli *et al*., 2000; Gallagher *et al*., 2000; Vogeley *et al*., 2001), only Baron-Cohen *et al*. (1999*b*) found amygdala activation. In the Baron-Cohen *et al*. (1999*b*) experiment the participants were required to read a mental or emotional state from an individual's eyes; therefore, their finding may reflect the amygdala's responsiveness to eye gaze information (Kawashima *et al*., 1999). This suggests that while the amygdala may yet prove to be involved in the development of ToM (Fine *et al*., 2001), it is perhaps not involved in the on-line representation of the mental states of others.

Conclusion

As far as we are aware, this is the first study to investigate the neural systems involved in the response to intentional and unintentional social norm transgressions. Consistent with *a priori* hypotheses, we found that the neural response to intentional and unintentional violations of social norms involved systems previously found to be implicated in the representation of the mental states of others; namely, medial prefrontal and temporal areas. In addition, and also as predicted, the neural response to intentional and unintentional violations of social norms involved systems that respond to the aversive emotional reactions of others, in particular others' anger; namely, the lateral orbitofrontal (BA 47) and medial prefrontal cortices. Interestingly, the response was very similar for both social norm violations and embarrassing conditions, albeit stronger for the norm violations. This suggests that a similar computational process, involving the representation of the mental states of others as well as expectations of social disapproval, is implemented when processing either an intentional social norm violation or an embarrassing situation. The greater activity of the regions associated with ToM by the intentional social norm violations relative to the embarrassing conditions may reflect attempts by the system to determine the protagonist's intention for engaging in the norm violation. Clearly, the present findings have direct implications for understanding the pathology of patients who exhibit social behavioural problems associated with frontal lobe lesions. In addition, they may provide insight into the social problems faced by some patients with frontotemporal dementia.

Acknowledgements

The authors wish to thank H. Gallagher for technical support. This work was supported by a program grant to R.J.D. from the Wellcome Trust and Medical Research Council and Department of Health (VISPED initiative) support of R.J.R.B. S.B. was funded by the Fyssen Foundation. J.L.A. is supported by the Brain Research Trust.

References

Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR. Impairment of social and moral behavior related to early damage in human prefrontal cortex. Nat Neurosci 1999; 2: 1032–7.

Averill JR. Anger and aggression: an essay on emotion. New York: Springer-Verlag; 1982.

Baron-Cohen S. Mindblindness: an essay on autism and theory of mind. Cambridge (MA): MIT Press; 1995.

Baron-Cohen S, O'Riordan M, Stone V, Jones R, Plaisted K.

Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. J Autism Dev Disord 1999a; 29: 407–18.

Baron-Cohen S, Ring HA, Wheelwright S, Bullmore ET, Brammer MJ, Simmons A, et al. Social intelligence in the normal and autistic brain: an fMRI study. Eur J Neurosci 1999b; 11: 1891–8.

Baron-Cohen S, Ring HA, Bullmore ET, Wheelwright S, Ashwin C, Williams SC. The amygdala theory of autism. [Review]. Neurosci Biobehav Rev 2000; 24: 355–64.

Bechara A, Damasio H, Damasio AR. Emotion, decision making and the orbitofrontal cortex. [Review]. Cereb Cortex 2000a; 10: 295–307.

Bechara A, Tranel D, Damasio H. Characterization of the decision-making deficit of patients with ventromedial prefrontal cortex lesions. Brain 2000b; 123: 2189–202.

Blair RJ. Neurocognitive models of aggression, the antisocial personality disorders and psychopathy. [Review]. J Neurol Neurosurg Psychiatry 2001; 71: 727–31.

Blair RJ, Cipolotti L. Impaired social response reversal: a case of 'acquired sociopathy'. Brain 2000; 123: 1122–41.

Blair RJ, Curran HV. Selective impairment in the recognition of anger induced by diazepam. Psychopharmacologia 1999; 147: 335–8.

Blair RJ, Morris JS, Frith CD, Perrett DI, Dolan RJ. Dissociable neural responses to facial expressions of sadness and anger. Brain 1999; 122: 883–93.

Blair RJ, Colledge E, Murray L, Mitchell DG. A selective impairment in the processing of sad and fearful expressions in chilldren with psychopathic tendencies. J Abnorm Child Psychol 2001; 29: 491–8.

Blair RJ, Frith U, Smith N, Abell F, Cipolotti L. Fractionation of visual memory: agency detection and its impairment in autism. Neuropsychologia 2002; 40: 108–18.

Blumer D, Benson DF. Personality changes with frontal and temporal lobe lesions. In: Benson DF, Blumer D, editors. Psychiatric aspects of neurological disease. New York (NY): Grune and Stratton; 1975. p. 151–70.

Bonda E, Petrides M, Ostry D, Evans A. Specific involvement of human parietal systems and the amygdala in the perception of biological motion. J Neurosci 1996; 16: 3737–44.

Borrill JA, Rosen BK, Summerfield AB. The influence of alcohol on judgement of facial expression of emotion. Br J Med Psychol 1987; 60: 71–7.

Brunet E, Sarfati Y, Hardy-Bayle MC, Decety J. A PET investigation of the attribution of intentions with a nonverbal task. Neuroimage 2000; 11: 157–66.

Burgess PW, Wood RL. Neuropsychology of behavior disorders following brain injury. In: Wood RL, editor. Neurobehavioural sequelae of traumatic brain injury. London: Taylor and Francis; 1990. p. 110–33.

Bush G, Luu P, Posner MI. Cognitive and emotional influences in anterior cingulate cortex. Trends Cogn Sci 2000; 4: 215–22.

Carmichael ST, Price JL. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. J Comp Neurol 1995; 363: 642–64.

Castelli F, Happe F, Frith U, Frith C. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. Neuroimage 2000; 12: 314–25.

Chao LL, Martin A, Haxby JV. Are face-responsive regions selective only for faces? Neuroreport 1999; 10: 2945–50.

Cipolotti L, Robinson G, Blair J, Frith U. Fractionation of visual memory: evidence from a case with multiple neurodevelopmental impairments. Neuropsychologia 1999; 37: 455–65.

Coull JT, Middleton HC, Robbins TW, Sahakian BJ. Clonidine and diazepam have differential effects on tests of attention and learning. Psychopharmacology (Berl) 1995; 120: 322–32.

Critchley HD, Elliott R, Mathias CJ, Dolan RJ. Neural activity relating to generation and representation of galvanic skin conductance responses: a functional magnetic resonance imaging study. J Neurosci 2000; 20: 3033–40.

Critchley HD, Melmed RN, Featherstone E, Mathias CJ, Dolan RJ. Brain activity during biofeedback relaxation: a functional neuroimaging investigation. Brain 2001; 124: 1003–12.

Damasio AR. Descartes' error: emotion, reason and the human brain. New York (NY): G. P. Putnam; 1994.

Davidson RJ, Putnam KM, Larson CL. Dysfunction in the neural circuitry of emotion regulation—a possible prelude to violence. [Review]. Science 2000; 289: 591–4.

Dewey M. Living with Asperger's syndrome. In: Frith U, editor. Autism and Asperger's syndrome. Cambridge (UK): Cambridge University Press; 1991. p. 184–206.

Dias R, Robbins TW, Roberts AC. Dissociation in prefrontal cortex of affective and attentional shifts. Nature 1996; 380: 69–72.

Dougherty DD, Shin LM, Alpert NM, Pitman RK, Orr SP, Lasko M, et al. Anger in healthy men: a PET study using script-driven imagery. Biol Psychiatry 1999; 46: 466–72.

Eslinger PJ. Neurological and neuropsychological bases of empathy. [Review]. Eur Neurol 1998; 39: 193–9.

Evans AC, Kamber M, Collins DL, Macdonald D. An MRI-based probabilistic atlas of neuroanatomy. In: Shorvon S, Fish D, Andermann F, Bydder GM, Stefan H, editors. Magnetic resonance scanning and epilepsy. New York (NY): Plenum Press; 1994. p. 263–74.

Farrow TF, Zheng Y, Wilkinson ID, Spence SA, Deakin JF, Tarrier N, et al. Investigating the functional anatomy of empathy and forgiveness. Neuroreport 2001; 12: 2433–8.

Fine C, Lumsden J, Blair RJ. Dissociation between 'theory of mind' and executive functions in a patient with early left amygdala damage. Brain 2001; 124: 287–98.

Fletcher PC, Happe F, Frith U, Baker SC, Dolan RJ, Frackowiak RS, et al. Other minds in the brain: a functional imaging study of 'theory of mind' in story comprehension. Cognition 1995; 57: 109–28.

Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC.

Assessing the significance of focal activations using their spatial extent. Hum Brain Mapp 1994; 1: 210–20.

Frith CD, Frith U. Interacting minds – a biological basis. [Review]. Science 1999; 286: 1692–5.

Frith U. Autism: explaining the enigma. Oxford: Blackwell; 1989.

Gallagher HL, Happe F, Brunswick N, Fletcher PC, Frith U, Frith CD. Reading the mind in cartoons and stories: an fMRI study of 'theory of mind' in verbal and nonverbal tasks. Neuropsychologia 2000; 38: 11–21.

Garland H, Brown BR. Face-saving as affected by subjects' sex, audiences' sex and audience expertise. Sociometry 1972; 35: 280–9.

Goel V, Grafman J, Sadato N, Hallett M. Modeling other minds. Neuroreport 1995; 6: 1741–6.

Grafman J, Schwab K, Warden D, Pridgen BS, Brown HR, Salazar AM. Frontal lobe injuries, violence, and aggression: a report of the Vietnam Head Injury Study. Neurology 1996; 46: 1231–8.

Grattan LM, Bloomer RH, Archambault FX, Eslinger PJ. Cognitive flexibility and empathy after frontal lobe lesion. Neuropsychiatry Neuropsychol Behav Neurol 1994; 7: 251–9.

Gregory CA, Hodges JR. Clinical features of frontal lobe dementia in comparison to Alzheimer's disease. [Review]. J Neural Transm Suppl 1996; 47: 103–23.

Grezes J, Costes N, Decety J. The effects of learning and intention on the neural network involved in the perception of meaningless actions. Brain 1999; 122: 1875–87.

Happe F, Malhi GS, Checkley S. Acquired mind-blindness following frontal lobe surgery? A single case study of impaired 'theory of mind' in a patient treated with stereotactic anterior capsulotomy. Neuropsychologia 2001; 39: 83–90.

Hare RD, Quinn MJ. Psychopathy and autonomic conditioning. J Abnorm Psychol 1971; 77: 223–35.

Harmer CJ, Bhagwagar Z, Cowen PJ, Goodwin GW. Acute adiminstration of citalopram in healthy volunteers facilitates recognition of happiness and fear. J Psychopharmacol 2001a; 15 Suppl: A16.

Harmer CJ, Thilo KV, Rothwell JC, Goodwin GM. Transcranial magnetic stimulation of medial-frontal cortex impairs the processing of angry facial expressions. Nat Neurosci 2001b; 4: 17–8.

Hecaen H, Albert ML. Human neuropsychology. New York (NY): John Wiley; 1978.

Hornak J, Rolls ET, Wade D. Face and voice expression identification in patients with emotional and behavioural changes following ventral frontal lobe damage. Neuropsychologia 1996; 34: 247–61.

Kawashima R, Sugiura M, Kato T, Nakamura A, Hatano K, Ito K, et al. The human amygdala plays an important role in gaze monitoring: a PET study. Brain 1999; 122: 779–83.

Kesler-West ML, Andersen AH, Smith CD, Avison MJ, Davis CE, Kryscio RJ, et al. Neural substrates of facial emotion processing using fMRI. Brain Res Cogn Brain Res 2001; 11: 213–26.

Levenston GK, Patrick CJ, Bradley MM, Lang PJ. The psychopath as observer: emotion and attention in picture processing. J Abnorm Psychol 2000; 109: 373–85.

Lishman WA. Brain damage in relation to psychiatric disability after head injury. Br J Psychiatry 1968; 114: 373–410.

Lough S, Gregory C, Hodges JR. Dissociation of social cognition and executive function in frontal variant frontotemporal dementia. Neurocase 2001; 7: 123–30.

Lund and Manchester Groups. Clinical and neuropathological criteria for frontotemporal dementia. [Review]. J Neurol Neurosurg Psychiatry 1994; 57: 416–8.

Miller BL, Darby A, Benson DF, Cummings JL, Miller MH. Aggressive, socially disruptive and antisocial behaviour associated with fronto-temporal dementia. Br J Psychiatry 1997; 170: 150–4.

Mitchell DGV, Colledge E, Leonard A, Blair RJR. Risky decisions and response reversal: is there evidence of orbitofrontal cortex dysfunction in psychopathic individuals? Neuropsychologia. In press 2002.

Neary D, Snowden JS, Gustafson L, Passant U, Stuss D, Black S, et al. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. [Review]. Neurology 1998; 51: 1546–54.

Oram MW, Perrett DI. Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. J Neurophysiol 1996; 76: 109–29.

Panksepp J. Affective neuroscience: the foundations of human and animal emotions. New York (NY): Oxford University Press; 1998.

Paus T, Koski L, Caramanos Z, Westbury C. Regional differences in the effects of task difficulty and motor output on blood flow response in the human anterior cingulate cortex: a review of 107 PET activation studies. [Review]. Neuroreport 1998; 9: R37–47.

Pennington BF, Bennetto L. Main effects or transaction in the neuropsychology of conduct disorder? Commentary on: The neuropsychology of conduct disorder. Dev Psychopathol 1993; 5: 153–64.

Perrett DI, Smith PA, Mistlin AJ, Chitty AJ, Head AS, Potter DD, et al. Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: a preliminary report. Behav Brain Res 1985; 16: 153–70.

Pietrini P, Guazzelli M, Basso G, Jaffe K, Grafman J. Neural correlates of imaginal aggressive behavior assessed by positron emission tomography in healthy subjects. Am J Psychiatry 2000; 157: 1772–81.

Price BH, Daffner KR, Stowe RM, Mesulam MM. The comportmental learning disabilities of early frontal lobe damage. Brain 1990; 113: 1383–93.

Puce A, Allison T, Bentin S, Gore JC, McCarthy G. Temporal cortex activation in humans viewing eye and mouth movements. [Review]. J Neurosci 1998; 18: 2188–99.

Raine A, Lencz T, Bihrle S, LaCasse L, Colletti P. Reduced prefrontal gray matter volume and reduced autonomic activity in antisocial personality disorder. Arch Gen Psychiatry 2000; 57: 119–27.

Rogers RD, Blackshaw AJ, Middleton HC, Matthews K, Hawtin K, Crowley C, et al. Tryptophan depletion impairs stimulus–reward learning while methylphenidate disrupts attentional control in healthy young adults: implications for the monoaminergic basis of impulsive behaviour. Psychopharmacologia 1999; 146: 482–91.

Rolls ET. The orbitofrontal cortex and reward. [Review]. Cereb Cortex 2000; 10: 284–94.

Rowe AD, Bullock PR, Polkey CE, Morris RG. 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. Brain 2001; 124: 600–16.

Saver JL, Damasio AR. Preserved access and processing of social knowledge in a patient with acquired sociopathy due to ventromedial frontal damage. Neuropsychologia 1991; 29: 1241–9.

Semin GR, Manstead AS. The social implications of embarrassment displays and restitution behavior. Eur J Soc Psychol 1982; 12: 367–77.

Sprengelmeyer R, Rausch M, Eysel UT, Przuntek H. Neural structures associated with the recognition of facial expressions of basic emotions. Proc R Soc Lond B Biol Sci 1998; 265: 1927–31.

Stevens D, Charman T, Blair RJ. Recognition of emotion in facial expressions and vocal tones in children with psychopathic tendencies. J Genet Psychol 2001; 162: 201–11.

Stone VE, Baron-Cohen S, Knight RT. Frontal lobe contributions to theory of mind. J Cogn Neurosci 1998; 10: 640–56.

Stuss DT, Benson DF. The frontal lobes. New York (NY): Raven Press; 1986.

Stuss DT, Gow CA, Hetherington CR. 'No longer Gage': frontal lobe dysfunction and emotional changes. [Review]. J Consult Clin Psychol 1992; 60: 349–59.

Stuss DT, Gallup GG Jr, Alexander MP. The frontal lobes are necessary for 'theory of mind'. Brain 2001; 124: 279–86.

Vogeley K, Bussfeld P, Newen A, Herrmann S, Happe F, Falkai P, et al. Mind reading: neural mechanisms of theory of mind and self-perspective. Neuroimage 2001; 14: 170–81.

Volavka J. Neurobiology of violence. Washington (DC): American Psychiatric Press; 1995.

Volkow ND, Tancredi L. Neural substrates of violent behaviour. A preliminary study with positron emission tomography. Br J Psychiatry 1987; 151: 668–73.

Worsley KJ, Friston KJ. Analysis of fMRI time-series revisited—again [letter]. Neuroimage 1995; 2: 173–81.

Received November 12, 2001. Revised February 14, 2002. Accepted February 18, 2002