Impaired social response reversal
A case of 'acquired sociopathy'

R. J. R. Blair[1,2] and L. Cipolotti[3]

[1]*Institute of Cognitive Neuroscience and* [2]*Department of Psychology, University College London and* [3]*Department of Clinical Neuropsychology, National Hospital for Neurology and Neurosurgery, London, UK*

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

Summary
In this study, we report a patient (J.S.) who, following trauma to the right frontal region, including the orbitofrontal cortex, presented with 'acquired sociopathy'. His behaviour was notably aberrant and marked by high levels of aggression and a callous disregard for others. A series of experimental investigations were conducted to address the cognitive dysfunction that might underpin his profoundly aberrant behaviour. His performance was contrasted with that of a second patient (C.L.A.), who also presented with a grave dysexecutive syndrome but no socially aberrant behaviour, and five inmates of Wormwood Scrubs prison with developmental psychopathy. While J.S. showed no reversal learning impairment, he presented with severe difficulty in emotional expression recognition, autonomic responding and social cognition. Unlike the comparison populations, J.S. showed impairment in: the recognition of, and autonomic responding to, angry and disgusted expressions; attributing the emotions of fear, anger and embarrassment to story protagonists; and the identification of violations of social behaviour. The findings are discussed with reference to models regarding the role of the orbitofrontal cortex in the control of aggression. It is suggested that J.S.'s impairment is due to a reduced ability to generate expectations of others' negative emotional reactions, in particular anger. In healthy individuals, these representations act to suppress behaviour that is inappropriate in specific social contexts. Moreover, it is proposed that the orbitofrontal cortex may be implicated specifically either in the generation of these expectations or the use of these expectations to suppress inappropriate behaviour.

Introduction

Frontal lobe damage has long been linked to emotional/personality changes such as euphoria, irresponsibility, lack of affect and lack of concern for the present or future (e.g. Hecaen and Albert, 1978; Stuss and Benson, 1986). Patients with aberrant behaviour following frontal lesions have been reported to show impairments in emotional expression recognition (Hornak *et al.*, 1996) and perform poorly on self-report measures of empathy (e.g. 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 both when the lesions are acquired early in life and when they occur in adulthood (e.g. Burgess and Wood, 1990; Price *et al.*, 1990; Pennington and Bennetto, 1993; Grafman *et al.*, 1996). For example, Burgess and Wood reported that 'Following serious [frontal] brain injury, many patients find it difficult to tolerate frustration and react impulsively to minor forms of pressure or provocation' (Burgess and Wood, 1990, p. 122). The aggression is sporadic, stress-related and directed more frequently against physical objects than against other individuals. Within the frontal lobes, it is orbitofrontal (ventral) and medial frontal damage that has been linked primarily to these emotional and social behavioural changes (e.g. Damasio, 1994; Volavka, 1995; Grafman *et al.*, 1996).

Probably the best-documented case of a patient with

emotional and behavioural changes following lesions to the orbitofrontal cortex is the patient E.V.R. (e.g. Eslinger and Damasio, 1985; Damasio *et al.*, 1990, 1991; Saver and Damasio, 1991; Bechara *et al.*, 1994, 1997; Damasio, 1994). Prior to his orbitofrontal meningioma, E.V.R. was a successful professional and was happily married. However, subsequent to his surgery, his social conduct was profoundly affected. He was divorced twice and entered into disastrous business ventures that led to predictable bankruptcy. This case prompted Damasio's introduction of the term 'acquired sociopathy'. With this, he aimed to characterize individuals who, following acquired lesions of the orbitofrontal cortex, fulfilled the DSM-III (Diagnostic and Statistical Manual of mental disorders) diagnostic criteria for 'sociopathic disorder'. These diagnostic criteria include being 'reckless regarding others' personal safety', 'lack of remorse' and 'defective planning' (DSM-III, 1980). Moreover, Damasio and colleagues proposed that psychopathy might be the developmental counterpart of the clinical picture presented by such patients (Damasio *et al.*, 1990). Psychopathy is characterized by high levels of aggression and antisocial behaviour performed without guilt or empathy for the victim (Hare, 1991).

Damasio proposed that impairment in a 'somatic marker system' caused the aberrant social behaviour (e.g. Damasio, 1994). This system tags internal representations with a somatic (body state) marker, i.e. an autonomic nervous system response. These somatic markers aid decision making: they 'force attention on the negative outcome to which a given action may lead, and function as an automated alarm signal that says: Beware of danger ahead if you choose the option that leads to this outcome. The signal may lead you to reject, *immediately*, the negative course of action and thus make you choose among other alternatives. The automated signal protects you against future losses, without further ado, and then allows you *to choose among other alternatives*' (Damasio, 1994, p. 173, italics in original).

Two major findings are associated with the somatic marker hypothesis. First, patients with lesions to the ventromedial frontal cortex fail to show autonomic responses to visually presented social stimuli (scenes of social disaster, mutilation and nudity) under passive viewing conditions (Damasio *et al.*, 1990, 1991). Passive viewing conditions require the participant only to look at the pictures of scenes. It should be noted that these patients did, however, show appropriate autonomic responses to the same stimuli under active viewing conditions. These involved attention-inducing instructions such as being asked to describe the pictures. Secondly, patients with lesions to the ventromedial frontal cortex perform poorly on the Four-Pack Card-Playing task (e.g. Bechara *et al.*, 1994, 1997). In this task, participants are presented with four packs of cards. Two packs of cards result in high reward but even higher punishments and, if played continuously, result in a net loss. The two other packs result in low reward but even lower punishments and, if played continuously, result in a net gain. The participants have to

learn to avoid the high reward, net loss packs in favour of the low reward, net gain packs. Healthy participants learned to take from the low reward packs and showed skin conductance responses (i.e. warning somatic markers) before the selection of a card from the disadvantageous packs. In contrast, patients with ventromedial damage continued to choose from the disadvantageous packs and failed to show skin conductance responses before their choices from these packs.

Four other major proposals have been offered to account for the aberrant emotional and social behaviour of patients with frontal lobe lesions. Rolls proposed that the inappropriate social behaviour shown by patients such as E.V.R. may 'be related to a dysfunction in altering behaviour appropriately in response to a change in reinforcement contingencies' (Rolls, 1996, p. 1438). He suggested that orbitofrontal cortex is associated with 'rapid stimulus–reinforcement association learning, and the correction of these associations when reinforcement contingencies in the environment change' (Rolls, 1996, p. 1437; see also Dias *et al.*, 1996; LeDoux, 1998). He argued that 'the orbitofrontal cortex is involved in emotional responses by correcting stimulus–reinforcement associations when they become inappropriate' (Rolls, 1996, p. 1438). Rolls and colleagues investigated this hypothesis using a visual discrimination task that required the patient to learn that touching one stimulus when it appeared on a computer screen resulted in reward while touching another visual stimulus resulted in punishment (Rolls *et al.*, 1994). After the patients learnt the visual discrimination, the reinforcement contingencies were reversed unexpectedly. Patients with lesions to the orbitofrontal cortex made more errors following contingency reversal, and completed fewer reversals than control populations with either more dorsolateral frontal or posterior lesions. Moreover, the patient's task performance correlated inversely with their level of disinhibited/socially inappropriate behaviour, as indexed by care-giver report.

Grafman interpreted the patient's impairment in terms of an inability to access 'social schema knowledge' stored in the frontal lobes (e.g. Grafman, 1994; Grafman *et al.*, 1996). Social schema knowledge is thought to inhibit aberrant behaviour. Patients with orbitofrontal cortex lesions who cannot access social schema knowledge fail to inhibit aberrant behaviour, such as physical threats and aggression. However, this position has received relatively little empirical testing. Dimitrov and colleagues found that some patients with frontal lobe lesions showed atypical performance on a task investigating ability to rate solutions to social problems according to their effectiveness (Dimitrov *et al.*, 1996). However, the patients who showed atypical performance were also those with the lowest IQs and poorest performance on the Wisconsin Card Sorting Test (WCST). It is thus unclear whether generalized intellectual or executive difficulties might explain their impaired performance on the experimental task. Moreover, Saver and Damasio (Saver and Damasio, 1991) found that, contrary to Grafman's hypothesis, E.V.R. showed intact social knowledge on tasks such as the

cartoon predictions test (O'Sullivan and Guilford, 1976) and the moral judgement interview (Colby and Kohlberg, 1987).

Baron-Cohen interpreted the patient's impairment in terms of damage to the neural circuit mediating Theory of Mind (Baron-Cohen, 1995). Theory of Mind is defined as the ability to represent the mental states of other individuals. The suggestion is that individuals who, because of orbitofrontal damage, cannot represent the mental states of others will be unable to modulate their emotional and social behaviour. In line with this, Price and colleagues reported two cases who performed poorly on a test of visual perspective taking (Price *et al.*, 1990), and a SPECT neuroimaging study reported that the orbitofrontal cortex was implicated in Theory of Mind processing (Baron-Cohen *et al.*, 1994). However, it is unclear whether visual perspective taking requires the representation of mental states (Leslie and Frith, 1988). However, two other functional imaging studies investigating the neural substrate of Theory of Mind reported medial frontal, rather than orbitofrontal, cortex involvement (Fletcher *et al.*, 1995; Goel *et al.*, 1995).

Brothers (Brothers, 1995, 1997) interpreted her patients' impairment in terms of a social 'editor'. This editor is 'specialised for processing others' social intentions . . .' (Brothers, 1997, p. 27). It attributes others' intentions by responding to significant gestures and expressions. The editor 'encourages the rest of the brain to report on features of the social environment' (Brothers, 1997, p. 15); effectively, it focuses attention on stimuli, such as faces and expressions, that are particularly important in social interactions. The editor is considered a unitary system 'specialised for responding to social signals of all kinds, a system that would ultimately construct representations of the mind' (Brothers, 1997, p. 27). She specifically argues against the idea of dissociable systems for what she terms 'cold' social cognition (i.e. the attribution of belief; Theory of Mind) and 'hot' social cognition (i.e. processing the emotional expressions of others).

Here we report a patient, J.S., who, following a head injury, developed a marked dysexecutive syndrome, characterized by 'acquired sociopathy'. Our aim was to investigate which dysfunctional cognitive systems underpin his profound emotional and social behavioural impairment. We compared his performance with that of another patient with grave dysexecutive syndrome, C.L.A., who did not present with 'acquired sociopathy'. Our aim was to disentangle those executive systems that may be crucial for social cognition from those that are crucial for general executive processing. We also compared the performance of J.S. with a small population of patients with developmental psychopathy. Our aim was to determine the commonalties and differences between 'acquired sociopathy' and developmental psychopathy.

Case report

J.S. is a 56-year-old, right-handed man who worked as an electrical engineer. In November 1996, he was admitted to

an Accident and Emergency Department after he was found collapsed and unconscious with evidence of trauma to the right frontal region. There was some suggestion of a witnessed generalized convulsion, but no other history was available. On admission, the patient was confused, with a reduced level of consciousness. On examination, he presented with a Glasgow coma scale of 9. There were disconjuntate eye movements with divergent squint but no nystagmus, flaccid tone with increased reflexes and down going plantars. He was first admitted to the Intensive Therapy Unit and then transferred to the neurological ward where he stayed until February 1997. During this time, he was noted to present with memory loss and a 'dysexecutive syndrome'. In particular, it was noted that his behaviour was gravely disturbed. He behaved bizarrely, for example riding around on a hospital trolley (gurney), he was uncooperative and he had several unpredictable episodes of aggression. An EEG showed irregular slowing which had a maximum anteriorly. An enhanced CT scan showed low-density abnormalities in both frontal lobes involving the orbitofrontal cortex. In addition, there was involvement of the left temporal lobe, almost certainly comprising the left amygdala because the uncus was of a slightly lower density on the left, and the left temporal horn was not visible (see Fig. 1).

Subsequently, he was transferred to a long-term rehabilitation hospital. During his stay, he assaulted and wounded a member of staff, frequently threw objects and furniture at people and was aggressive towards other patients. Staff and patients felt at risk and frightened by his presence.

In July 1997, he was admitted to the National Hospital for Neurology and Neurosurgery for further evaluation of his grave behavioural disturbances. An MRI scan was performed; however, its results were confounded by movement artefacts. During his admission, it was established that he had no psychiatric history prior to his accident. He was described by a relative as being premorbidly a quiet, rather withdrawn person who was never aggressive. Following discharge from the National Hospital, he was re-admitted to a long-term rehabilitation hospital.

Behavioural assessment

Since the injury, J.S. fulfilled the DSM-IV criteria for Antisocial Personality Disorder (DSM-IV, 1994). J.S. 'failed to conform to social norms' and was notably 'irritable and aggressive'. His episodes of property damage and violence were frequent and were elicited after little provocation; e.g. an alteration in routine. He was 'reckless regarding others' personal safety'; on one occasion he continued to push around a wheelchair-bound patient despite her screams of terror. His 'lack of remorse' was striking; he never expressed any regrets about the nurses he hit. He failed to accept responsibility for his actions, justifying his violent episodes in terms of the failures of others (e.g. they were too slow). He frequently 'failed to plan ahead', leaving the hospital regularly to wander about London with 'no clear goal for the period of travel or
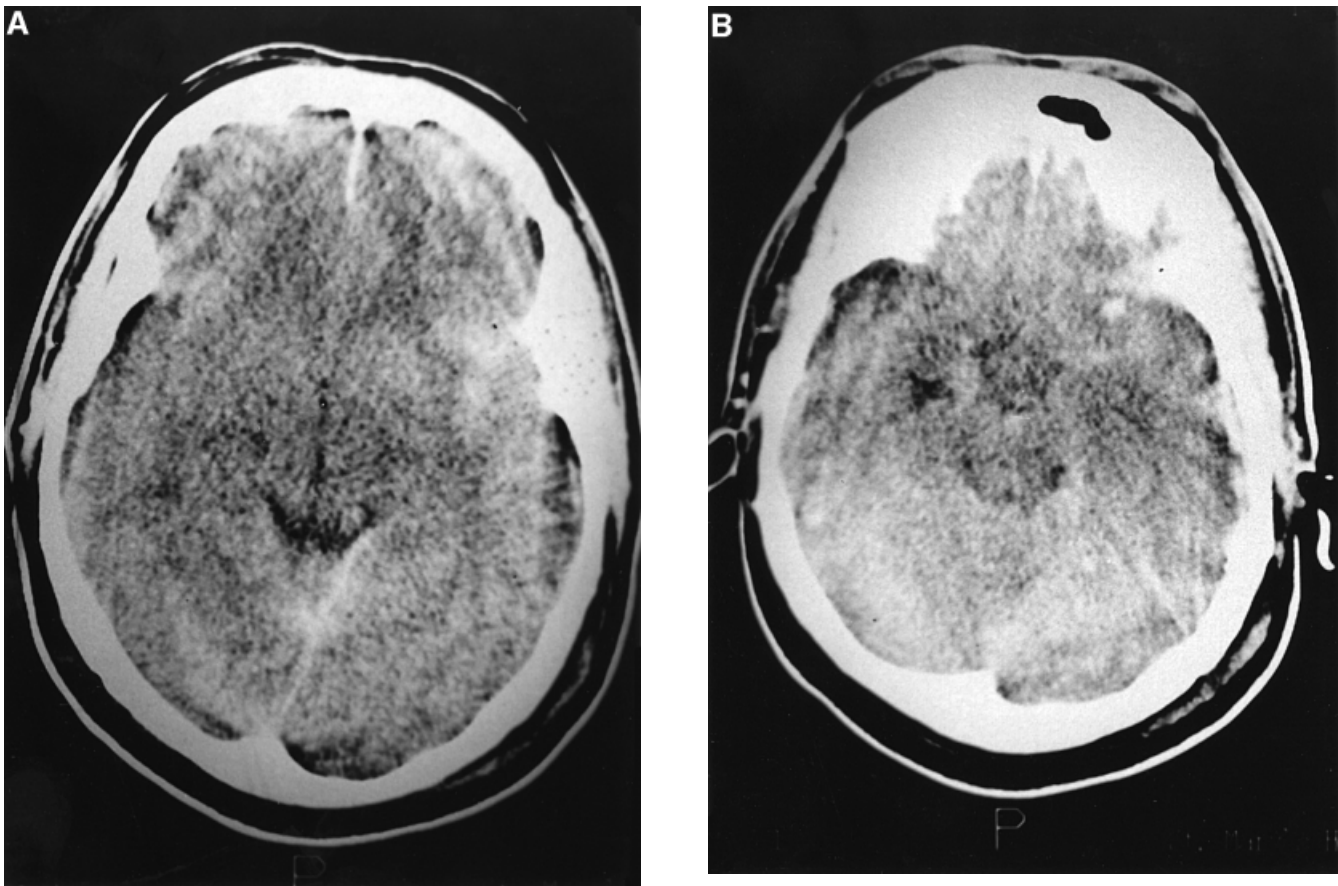
Fig. 1 (**A**) Enhanced CT scan image showing low density abnormalities bilaterally frontally. Orbitofrontal cortex was damaged on both sides. (**B**) Enhanced CT scan image showing low density abnormalities in the left temporal lobe. On the left side, the uncus was of a slightly lower density and the left horn was not visible. This was taken to indicate that the left amygdala was also abnormal.

clear idea about when the travel will terminate'. He showed 'an inability to sustain consistent work behaviour'. Since the accident, he could not hold employment due to his interpersonal difficulties. In summary, J.S. fulfilled the criteria for acquired sociopathy except that he lacked premorbid aberrant behaviour.

While J.S. showed reduced 'empathy' and profoundly disturbed social interactions, there were no indications of self-neglect, changes in eating habits, hyper-sexuality or oral exploratory behaviour. In summary, J.S. showed no indications of the Klüver–Bucy syndrome.

Neuropsychological assessment
J.S. was assessed on a shortened form of the Wechsler Adult Intelligence Scale—Revised (WAIS-R) and obtained a verbal and a performance IQ in the average range (see Table 1). On Raven Advanced Matrices (Raven, 1965), he obtained a good average score. Reading performance on the National Adult Reading Test (NART) (Nelson and Willison, 1991) gave an estimated premorbid level of functioning in the superior range. On the recognition memory test (Warrington, 1984), his performance was in the high average range for verbal material and in the low average range for visual material. On

an additional visual recognition memory test, his performance was again in the low average range (Topographical Recognition Memory; Warrington, 1996). His nominal skills were impaired. His performance was poor on the Graded Difficulty Naming Test (McKenna and Warrington, 1980) and on the Oldfield picture-naming test (Oldfield and Wingfield, 1965). Few semantic errors were present in both tasks. J.S.'s single word comprehension was within the average range on a stringent Synonym Test (Warrington *et al.*, 1998). Visuo-perceptual and visuo-spatial skills, as assessed by two subtests of the Visual Object and Space Perception Battery, were within normal limits (Warrington and James, 1991).

Frontal executive functions
His performance on tests sensitive to frontal lobe damage was impaired (see Table 1). His responses on the Cognitive Estimates Test (Shallice and Evans, 1978) were bizarre (e.g. weight of a full pint of milk = '. . . 10 . . . no 7 pounds . . .'). On the WCST (Nelson, 1976), he obtained only three categories, making several perseverative errors, and the test was abandoned due to low frustration tolerance. On the Hayling Sentence Completion Test (Burgess and Shallice, 1996), he performed flawlessly but slowly on the response

Table 1 *Cognitive and frontal lobe test scores*

J.S.	Percentile	C.L.A.	Percentile	
Verbal IQ	98		96	
Performance IQ	92		–	
Raven's Advanced Matrices	8/12	50–75%ile	8/12	50–75%ile
NART IQ	123		117	
RM Words	48/50	>75%ile	42/50	10–25%ile
RM Faces	39/50	= 25%ile	31/50	= 25%ile
RM Topographical	21/30	= 25%ile	19/30	= 5%ile
GNT	7/30	<1%ile	27/30	= 95%ile
Oldfield	22/30			
Synonyms test	43/50	50–75%ile		
Object decision	17/20	>5% cut-off	19/20	>5% cut-off
Cube analysis	8/10	>5%ile	20/20	
Cognitive estimates	F		F	
Wisconsin	F		F	
Hayling Sentence Completion	F		F	
Proverbs	F		F	
Verbal Fluency	–		F	
Weigl Colour Form Sort	P		F	
Stroop Test	P		F	
Trail Making B	P		–	

NART = National Adult Reading Test; RM = recognition Memory; GNT = Graded Difficulty Naming Test; P/F = pass/fail; – = not tested.

initiation section. However, on the response suppression section, he was unable to produce words that were unrelated to the sentence. For example, for 'the captain wanted to stay with the sinking . . .' he could not suppress the response 'ship . . .'. On a proverb interpretation task, he gave a concrete interpretation for 6/10 proverbs. However, he had no difficulty with the Weigl Colour Form Sorting Test (Weigl, 1941), the Stroop test (Stroop, 1935) and the Trail-Making Test (Army Individual Test Battery, 1944). In addition, his speed of performance on the Trail-Making Test was normal.

Comment

J.S. presented with acquired sociopathy. His cognitive profile was characterized by marked frontal dysexecutive impairments and nominal difficulties. In the following experimental investigation, we explored the basis of J.S.'s acquired sociopathy.

Experimental investigation

The following experimental investigation was conducted over 1 week (July 17–24, 1997). This was introduced to J.S. as a way of understanding his behaviour more fully. He stated that he enjoyed participating as he found the ward boring. In order to investigate J.S.'s acquired sociopathy, reversal learning, expression recognition, emotional responding and social cognition tests were carried out.

Control subjects

J.S.'s performance was contrasted against that of a patient, C.L.A., who presented with a grave dysexecutive syndrome but no behavioural disturbance (see Appendix A for patient details) on tests of reversal learning, emotional expression processing and social cognition. For the reversal learning, expression recognition and autonomic responding and social cognition tasks, J.S.'s performance was contrasted with that of five male prison inmates with developmental psychopathy and five non-psychopathic control inmates held in the same institution (see Appendix B for population details). For the social cognition tasks, J.S.'s performance is also contrasted with 10 age- and IQ-matched normal participants (mean age = 54 years, SD = 7; mean NART IQ = 104, SD = 7).

Reversal learning tasks

Two reversal learning tasks were administered (the Four-Pack and the One-Pack Card-Playing tasks). These tasks were developed from Damasio's somatic marker hypothesis and Newman's response set modulation hypothesis (Newman *et al.*, 1987; Damasio, 1994). Both tasks assess the participant's ability to reverse a learnt response and are thus considered together below (see Dias *et al.*, 1996; Rolls, 1998)

Task 1: the Four-Pack Card-Playing task

Damasio and colleagues have reported that patients with acquired sociopathy are impaired on the Four-Pack Card-Playing task (Bechara *et al.*, 1994, 1997). We wished to determine whether J.S. would show impairment.

Procedure. This task used a version of the protocol described by Bechara and colleagues (Bechara *et al.*, 1994). In this task, participants are presented with four packs of

Table 2 *Reversal learning tasks: mean number of cards played in the four- and one-pack tests (standard deviations and ranges in parentheses)*

	J.S.	C.L.A.	Psychopathic inmates	Non-psychopathic inmates
Four-pack task				
A	11	13	14.2	12.8
Max = 100			(5.12, 10–23)	(2.05, 11–16)
B	13	82	23.6	19.4
Max = 100			(3.29, 20–28)	(6.19, 10–26)
C	36	3	32.4	36.2
Max = 100			(11.10, 18–45)	(5.26, 32–45)
D	40	2	29.8	31.6
Max = 100			(9.93, 17–44)	(7.23, 26–44)
One-pack task				
Max = 100	15	29	91.2	48.6
			(14.74, 66–100)	(32.84, 10–73)

cards. Two packs of cards result in high reward but even higher punishments and, if played continuously, result in a net loss. The two other packs result in low reward but even lower punishments and, if played continuously, result in a net gain. The participants have to learn to avoid the high reward, net loss packs in favour of the low reward, net gain packs.

Results. In Table 2, J.S.'s performance is contrasted with that of the comparison populations (C.L.A. and the psychopathic and non-psychopathic inmates). J.S. performed well within the range of the psychopathic and non-psychopathic inmates and the normal control participants reported by Bechara and colleagues (Bechara *et al.*, 1994). In contrast, C.L.A. performed poorly, perseverating in his choice of pack B (82% of his responses).

Comment. J.S.'s unimpaired performance on the Four-Pack Card-Playing task suggests that he did not present with somatic marker impairment. Moreover, given the task's reversal learning, J.S. showed unimpaired reversal learning.

Task 2: the One-Pack Card-Playing task
In Task 2, we wished to determine J.S.'s performance on a second reversal learning task that psychopathic individuals perform poorly (e.g. Newman *et al.*, 1987).

Procedure. This task used a version of the protocol previously described by Newman and colleagues (Newman *et al.*, 1987). In this task, the participant has to decide whether to play a card. Initially, the participant's choices to play are reinforced, but as the number of trials increases, the probability of reward decreases. The participant should terminate his responding before he receives greater levels of punishment than reward.

Results. In Table 2, the performance of J.S. is contrasted with that of the comparison populations (patient C.L.A. and

the psychopathic and non-psychopathic inmates). J.S. played few cards and stopped after the first couple of 'punishments'. Two of the non-psychopathic inmates responded similarly. C.L.A. and the remaining three non-psychopathic inmates played relatively few cards. In striking contrast, the psychopathic inmates played, on average, almost double the number of cards. Indeed, three continued playing the game until its end, resulting in the loss of all their tokens.

Comment. The results of Task 2 replicated and extended those of Task 1 indicating that J.S. had no reversal learning impairment. In other words, he could reverse his responses to stimuli that previously had been associated with reward when they became associated with punishment. This suggests that his behavioural difficulties cannot be attributed easily to disruption within those systems that are crucial for reversal learning.

Expression recognition and emotional responding tasks
Two tasks were designed to investigate J.S.'s processing of emotional stimuli. In the first, we investigated his ability to recognize and autonomically respond to human facial expressions (Task 3). In the second, we investigated his autonomic responses to more naturalistic stimuli, including human expressions and threatening objects (Task 4).

Task 3: facial expression processing
Patients with aberrant behaviour following orbitofrontal cortex lesions have been reported to show impairments in emotional expression recognition (Hornak *et al.*, 1996). We wished to determine whether J.S. would show similar deficits. In addition, we wished to determine whether any recognition deficit/impairment in skin conductance response (SCR) generation would be selective to particular emotion categories.

Table 3 *Performance on expression recognition and skin conductance responses to expression stimuli (standard deviations and ranges in parentheses)*

	J.S.	C.L.A.	Psychopathic inmates	Non-psychopathic inmates
Expression recognition score				
Surprise	16	20	18.0 (1.58, 16–20)	18.2 (1.30, 17–20)
Happiness	13*,†	17	19.6 (0.55, 19–20)	19.2 (0.84, 18–20)
Anger	9*,†,‡	15	18.6 (1.14, 17–20)	15.2 (1.64, 14–18)
Disgust	4*,†,‡	20	16.4 (3.21, 12–20)	16.6 (1.67, 14–18)
Sadness	9*,†	5	18.0 (1.58, 16–20)	18.2 (2.17, 15–20)
Fearfulness	5	6	9.8 (7.69, 2–18)	12.6 (3.44, 8–15)
SCRs to expression stimuli				
Surprise	0.204 (0.358, 0.0–0.458)	–	0.264 (0.102, 0.156–0.375)	0.316 (0.234, 0.136–0.677)
Happiness	0.214 (0.374, 0.0–0.778)	–	0.270 (0.125, 0.157–0.424)	0.225 (0.094, 0.126–0.378)
Disgust	0.080*,† (0.095, 0.0–0.336)	–	0.283 (0.188, 0.129–0.548)	0.412 (0.273, 0.193–0.845)
Anger	0.041*,† (0.055, 0.0–0.153)	–	0.363 (0.203, 0.213–0.639)	0.379 (0.199, 0.235–0.716)
Sadness	0.080† (0.172, 0.0–0.305)	–	0.146 (0.058, 0.101–0.229)	0.353 (0.138, 0.222–0.537)
Fearful	0.074† (0.118, 0.0–0.342)	–	0.116 (0.040, 0.081–0.172)	0.491 (0.300, 0.208–0.935)

*J.S.'s performance is significantly poorer than that of the lowest scoring psychopathic inmate; †J.S.'s performance is significantly poorer than that of the lowest scoring non-psychopathic inmate; ‡J.S.'s performance is significantly poorer than C.L.A.'s; – = not tested. χ^2 (d.f. = 1) analyses were used for all subject × expression recognition comparisons. One-way ANOVAs were used for all between-subject by SCR to expression comparisons.

Procedure. This task was based on the paradigm described by Calder and colleagues (Calder *et al.*, 1996). This method assesses recognition of six emotion expressions: surprise, happiness, anger, disgust, sadness and fearfulness. The stimuli are continuous tone images where two expressions have been morphed together.

Each face was presented to J.S. on a computer screen, with each stimulus subtending a horizontal visual angle of 3.6° and a vertical angle of 5.2°. There were six blocks of stimuli. In each block, the 30 stimuli were presented in a randomized order. The first block was counted as practice trials and the data obtained were not analysed. Each stimulus was presented for 3 s and there was a 4–6 s interval between each stimulus during which the screen was blank. J.S. was presented with a list of six response options (the six emotions) and instructed to name the displayed expression. In addition, his SCR was recorded in the 1–3 s interval following stimulus onset.

Results. In Table 3, the performance of J.S. on expression recognition is contrasted with that of patient C.L.A., and the psychopathic and non-psychopathic inmates. Chi-square

analyses were performed to explore J.S.'s ability relative to the comparison populations. In comparison with both inmate populations, J.S. was significantly impaired at recognizing happiness, anger, disgust and sadness, but not surprise and fearfulness. In comparison with C.L.A., J.S. was significantly impaired at recognizing anger and disgust, but not the other four expressions. Surprisingly, C.L.A. showed selective impairment in comparison with both inmate populations in sadness recognition.

In Table 3, the autonomic responses of J.S. to the six expressions are contrasted with those of the inmate populations. J.S. was profoundly impaired in generating autonomic responses to all emotional expressions other than surprise and happiness. ANOVAs (analyses of variance) revealed that J.S. produced significantly lower SCRs to anger and disgust expressions than both comparison groups. However, J.S. only produced significantly lower SCRs than the non-psychopathic inmates to the sad and fearful expressions. His SCRs to sad and fearful expressions were not significantly less than those of the psychopathic inmates.

Comment. In this task, J.S. was found to show severe

difficulties processing expression stimuli. Relative to C.L.A., who presented a selective impairment for the recognition of sadness, J.S. was impaired in the recognition of anger and disgust. Relative to the psychopathic and non-psychopathic inmates, J.S. was impaired in the recognition of anger, disgust, happiness and sadness. In addition, his SCRs to anger and disgust were lower than those shown by the psychopathic inmates. J.S.'s SCRs to sad and fearful expressions, as well as anger and disgust, were lower than those shown by the non-psychopathic inmates. The significance of these findings will be discussed.

Task 4: autonomic responses to environmentally salient visual stimuli

In Task 3, J.S. showed reduced autonomic responses to negative emotional expressions. In Task 4, we investigated J.S.'s autonomic responses to emotional expressions using naturalistic images that may be more powerful motivators of autonomic activity than the Ekman and Friesen-based stimuli used in Task 3 (Ekman and Friesen, 1975). In addition, we used other naturalistic stimuli (threatening and neutral objects) to explore whether his impaired autonomic responding was restricted to emotional expressions.

Procedure. This task is based on a paradigm described by Blair and colleagues (Blair *et al.*, 1997). The participant is presented with 50 stimuli in two blocks. Each block of 25 stimuli contains five sad, five angry and five neutral naturalistic expressions, five threatening objects (e.g. pointed weapons) and five neutral objects (e.g. household objects). The 25 stimuli within each block were presented in a random order. Moreover, block order was randomized across participants. There was a 6–8 s interval between each stimulus. SCRs were recorded in the 1–3 s following stimulus onset. In order to ensure that J.S. was attending to the stimuli, he was asked to give a score between 1 and 10 concerning how much he liked the picture.

Results. In Table 4, J.S.'s autonomic responses to the naturalistic stimuli are contrasted with five of the 10 healthy normal controls. In contrast to the comparison population, J.S. failed to show greater SCRs to the angry and sad naturalistic expressions than the neutral expressions [$F(1,18)$ < 1 for both comparisons]. He also failed to show greater SCRs to the threatening objects than the neutral objects [$F(1,18)$ < 0]. Comparisons with the healthy subject with the lowest SCRs to the stimuli revealed significant main effects for individual and significant individual × stimulus type interactions (sad/angry versus neutral face, or threatening object versus neutral object). Thus, the SCR difference in response to the sad and angry faces compared with the neutral stimulus was significantly greater in the comparison subject than in J.S. for angry versus neutral faces [$F(1,36) = 10.05$; $P < 0.05$], sad versus neutral faces [$F(1,36) = 12.63$;

Table 4 *Mean skin conductance responses to environmentally salient visual stimuli (standard deviations and ranges in parentheses)*

	J.S.	Healthy controls
Naturalistic expression stimuli		
Anger	0.058*	0.437
	(0.087, 0–0.550)	(0.198, 0.250–0.725)
Sadness	0.055*	0.421
	(0.135, 0–0.427)	(0.213, 0.246–0.790)
Neutral	0.042	0.134
	(0.087, 0–0.212)	(0.083, 0.070–0.273)
Objects		
Threatening	0.062*	0.561
	(0.138, 0–0.411)	(0.377, 0.245–1.20)
Neutral	0.109	0.251
	(0.185, 0–0.509)	(0.178, 0.076–0.500)

*J.S.'s performance is significantly poorer than that of the lowest responding healthy control. One-way ANOVAs were used for all between-subject by expression comparisons.

$P < 0.05$] and threatening objects versus neutral objects [$F(1,36) = 9.95$; $P < 0.01$].

Comment. The SCR data obtained in Task 4 replicate and extend the results found in Task 3. J.S. was hyporesponsive not only to the expression stimuli but also to threatening objects. Given J.S.'s profound deficits in Tasks 3 and 4, we wanted to be certain that they could not be attributed to some general impairment in face processing or the production of autonomic responses. Thus, we examined J.S.'s ability to process human faces (Task 5) and to generate SCRs to basic stimuli (Task 6).

Task 5: face processing tasks

To determine whether J.S. (and C.L.A.) could process facial stimuli adequately, we carried out the following series of tasks.

Famous Faces Naming Test. Twelve photographs of contemporary famous people were presented to J.S. (updated version of Warrington and James, 1967). J.S. (and C.L.A.) showed no impairment on this test; J.S. promptly named all 12 famous faces, while C.L.A. managed nine.

Perception of facial features in unfamiliar faces. Two tests focused on the perception of sex and age in facial features. The stimuli were photographs of unknown people and have been described previously (McNeil and Warrington, 1991).

Perception of sex. The stimuli were 20 photographs of unfamiliar faces in which the hair had been masked. The faces were presented individually and the patient was required to state the person's sex. J.S. and C.L.A. both performed flawlessly (score = 20/20 for both).

Perception of age. The test stimuli consisted of 24 sets of four unfamiliar faces. There were two conditions in which the patient was required to point to either the older or younger of the four. J.S. and C.L.A. both performed satisfactorily (score = 20/24 and 21/24, respectively).

Perception of facial configuration in unfamiliar faces. J.S.'s ability to identify faces from different angles and distances and under different lighting conditions was assessed using two face-matching tasks. C.L.A. was only presented with the first matching task.

Face matching. This test was adapted from De Renzi and colleagues (De Renzi *et al.*, 1969). It consisted of 12 pairs of photographs of female faces, each taken from a different view. Six pairs were of the same person and six of a different person. The patient stated whether the pair was of the same or different persons. J.S. and C.L.A. performed well on this task (score = 12/12 and 9/12, respectively).

Facial recognition test. Benton and Van Allen's facial recognition test was administered (Benton and Van Allen, 1968). J.S. obtained a score of 52, which is within the normal range.

Comment. J.S.'s (and C.L.A.'s) satisfactory performances indicated that both patients had normal face processing skills.

Task 6: autonomic responses to auditory stimuli

To determine J.S.'s ability to generate autonomic responses, we investigated his responsiveness to basic auditory stimuli such as a loud noise, his own name and an unknown name.

Procedure. In this task, three auditory stimuli were presented in six blocks. Within each block, the order of the three stimuli was randomized. The three stimuli types were a clap, the participant's name and the name of a person unknown to J.S. The SCR magnitude was recorded; defined as the greatest magnitude of departure from baseline occurring between 1 and 4 s after stimulus onset. The responses to the first four stimuli were considered practice trials and were not analysed.

Results. In Table 5, the performance of J.S. is contrasted with that of a subgroup of five of the 10 healthy normal controls. A series of one-way ANOVAs revealed that J.S.'s SCRs were not significantly lower than those shown by the controls for any of the three stimuli [$F(1,28)$ = 2.65, 2.16 and 0.89, n.s., for the clap, participant's name and unknown name, respectively]. A second series of one-way ANOVAs revealed that J.S. showed significantly larger responses to the clap and his own name than to the unknown name [$F(1,8)$ = 58.50, $P < 0.001$ and $F(1,8)$ = 3.94, $P < 0.05$, for the clap and own name, respectively]. The comparison participants performed similarly [$F(1,47)$ = 12.58, P < 0.01

Table 5 *Mean skin conductance responses to auditory stimuli (standard deviations and ranges in parentheses)*

	J.S.	Healthy controls
Clap	0.51	0.91
	(0.12, 0.40–0.66)	(1.05, 0.23–1.63)
Participant's name	0.19	0.48
	(0.10, 0.06–0.31)	(0.69, 0.16–0.99)
Unknown name	0.03	0.09
	(0.04, 0.0–0.14)	(0.15, 0.01–0.18)

One-way ANOVAs were used for all between-subject by expression comparisons. There were no significant differences between J.S.'s performance and that of the lowest performing healthy control.

and $F(1,47)$ = 5.47, $P < 0.05$ for the clap and own name, respectively].

Comment. J.S.'s performance indicated that he did not present with a basic inability to generate SCRs.

Social cognition tasks

The following tasks addressed four aspects of social cognition. First, J.S.'s ability to represent the emotions of others (Task 8). Secondly, his ability to represent the internal mental states of others (Task 9). Thirdly, his ability to process the appropriateness of behaviour in different social contexts (Tasks 10 and 11). Fourthly, his sensitivity to basic behavioural norms as they are embodied in the moral/conventional distinction test (Task 12).

Task 7: verbal comprehension

Because the social cognition tasks all involved story-type stimuli, we investigated J.S.'s ability to understand complex verbal material using the verbal comprehension subtest of the WAIS-R. J.S. obtained a high average score. This suggests normal verbal comprehension, and thus any impairment on the social cognition tasks could not be attributed to comprehension impairment.

Task 8: emotion attribution task

J.S.'s performance on Task 3 indicated that he was impaired in the recognition of emotional expressions. In Task 8, we investigated J.S.'s ability to attribute emotional states to others.

Procedure. This task used an expanded version of the paradigm described by Blair and colleagues (Blair *et al.*, 1995). In this task, the participant was presented with 67 short stories describing emotional situations and was asked what the main protagonists might feel in that situation. Fourteen stories were designed to elicit attributions of happiness, 14 of sadness, nine of fear, 10 of anger and 20

Table 6 *Number of correct answers on the emotion attribution task (standard deviations and ranges in parentheses)*

	J.S.	C.L.A.	Healthy controls	Psychopathic inmates	Non-psychopathic inmates
Happiness (Max = 14)	12	14	12.9 (0.99, 12–14)	11.8 (0.84, 11–13)	12.4 (1.14, 11–14)
Sadness (Max = 14)	12	14	11.6 (1.65, 9–14)	12.2 (1.48, 10–14)	11.4 (1.14, 11–13)
Fear (Max = 9)	1*,†,‡,¶	8	8.8 (0.42, 8–9)	8.4 (0.89, 7–9)	8.8 (0.45, 8–9)
Anger (Max = 10)	0*,†,‡,¶	10	7.2 (1.40, 5–10)	7.8 (0.84, 7–9)	8.2 (1.10, 8–10)
Embarrassment (Max = 20)	0*,†,‡,¶	9	13.5 (3.44, 7–18)	11.2 (4.21, 7–17)	12.6 (3.91, 8–18)

*J.S.'s performance is significantly poorer than that of the lowest scoring control; †J.S.'s performance is significantly poorer than that of the lowest scoring psychopathic inmate; ‡J.S.'s performance is significantly poorer than that of the lowest scoring non-psychopathic inmate; ¶J.S.'s performance is significantly poorer than C.L.A.'s. χ^2 analyses were used for all comparisons.

of embarrassment. Example stories are provided in Appendix C (i).

The task was scored according to the number of correct attributions made for each story category. Any positive affect (including relieved, overjoyed and ecstatic) was scored as correct for the happiness stories. Any sad affect (including sad, upset and devastated) was scored as correct for the sadness stories. Correct responses to the fearful stories included fearful, scared and petrified. Correct responses to the angry stories included angry, annoyed and irritated. Correct responses to the embarrassment stories included embarrassed, stupid and humiliated.

Results. In Table 6, the performance of J.S. on the emotion attribution task is contrasted with that of C.L.A., 10 healthy normal males, and five psychopathic and five non-psychopathic inmates. J.S. showed no difficulty making attributions of happiness and sadness to story characters. However, J.S. showed profound difficulty in attributing fear, anger or embarrassment. He ignored the threatening component in the fear stories, and stated that the protagonist would feel nothing or 'confusion' in the anger and embarrassment stories. C.L.A.'s performance did not differ significantly from that of the comparison populations.

Comment. The results of this task showed that J.S. had remarkable difficulty in attributing fear, anger and embarrassment to characters. The significance of these findings will be discussed further.

Task 9: advanced Theory of Mind task

The previous task indicated that J.S. was impaired in the attribution of fear, anger and embarrassment. In Task 9, we investigated whether J.S.'s emotion attribution impairment might reflect dysfunctional Theory of Mind.

Procedure. This task involves the participant reading 24 stories describing naturalistic social situations and being asked about why the characters behaved as they did (Happé, 1994). An example story is given in Appendix C (ii).

Three scores are generated from the participant's performance. The first, Total Score, indexes the participant's comprehension of the situation. The other two scores refer to the justifications the participant uses when interpreting the story characters' behaviour; i.e. reference to either the character's mental states or physical information. Example mental state and physical information justifications are given in Appendix C (ii).

Results. In Table 7, the performance of J.S. on the advanced Theory of Mind task is contrasted with that of C.L.A., and the psychopathic and non-psychopathic inmates. J.S. was unimpaired on this task. His comprehension performance (Total Score) was, like C.L.A.'s, flawless. Chi-square analyses demonstrated that J.S. was not significantly less likely to make mental state justifications than any of the comparison populations.

Comment. In this task, we documented that J.S. had no impairment in attributing mental states. This result suggests that his emotion attribution impairment was not underpinned by a Theory of Mind impairment.

Task 10: the moral/conventional distinction task

J.S. presented with profoundly aberrant behaviour. Grafman interpreted his findings of high levels of aggression in patients with lesions to the ventromedial frontal cortex in terms of an inability to access 'social schema knowledge' stored in the frontal lobes (e.g. Grafman, 1994; Grafman *et al.*, 1996). This might predict that J.S. should show impairment in social

Table 7 *Performance on the advanced theory of mind and moral/conventional distinction tasks (standard deviations and ranges in parentheses)*

	J.S.	C.L.A.	Psychopathic inmates	Non-psychopathic inmates
Performance on the advanced Theory of Mind task				
Total score	24	24	22.6	22.0
(Max = 24)			(1.95, 20–24)	(1.58, 20–24)
Justifications for correct	19	17	18.4	19.4
responses involving			(0.89, 18–20)	(1.67, 18–22)
mental states				
Physical information	5	7	4.2	2.6
			(1.79, 2–6)	(0.55, 2–3)
Performance on the moral/conventional distinction task				
Number of transgressions judged non-permissibility (Question 1, Max = 9)				
Moral 9	9	9.0	9.0	
Conventional	9	8	8.8 (0.45, 8–9)	8.6 (0.55, 8–9)
Total seriousness score (Question 2, Max = 90)				
Moral	90	71	87.2 (1.79, 85–90)	78.4 (6.84, 69–85)
Conventional	84	26	79.6 (3.51, 74–83)	29.8 (15.3, 17–54)
Number of transgressions judged non-permissible in the absence of rules (Questions 3 and 4, Max = 18)				
Moral	18	18	17.6 (0.55, 17–18)	18
Conventional	18	8	16.8 (1.10, 16–18)	7.4 (3.13, 4–10)

χ^2 (d.f. = 1) analyses were used for all analyses for the advanced Theory of Mind task. There were no significant differences between J.S.'s performance and those of the comparison populations.

rule knowledge. There are two different classes of social rules: moral (i.e. victim-based; hitting another individual) and conventional (i.e. social disorder-based; talking in class) which are distinguished by children from the age of 36 months (Smetana and Braeges, 1990). Moral transgressions are judged more serious than conventional transgressions, and they continue, unlike conventional transgressions, to be judged as non-permissible even in the absence of prohibiting rules. In Task 10, we investigated J.S.'s performance on the moral/conventional distinction test.

Procedure. The moral/conventional distinction stories were taken from the literature (e.g. Turiel *et al.*, 1987; Blair, 1995). Nine moral and nine conventional transgressions were presented to the participants in a randomized order. Four questions presented in a fixed order after each story had been read were used to assess the moral/conventional distinction. Examples of the transgressions used and the four questions are described in Appendix C (iii).

Results. In Table 7, the performance of J.S. on the moral/conventional distinction is contrasted with that of C.L.A., and the psychopathic and non-psychopathic inmates. J.S. showed no impairment in knowledge about the normal permissibility of the transgressions read to him. Unlike C.L.A. and the non-psychopathic subjects, J.S. judged both the moral and conventional transgressions as extremely serious. ANOVAs revealed that J.S. judged the conventional transgressions as significantly more serious than both C.L.A.

and the non-psychopathic inmates [$F(1,32) = 64.4$ and 18.0, 19.7, $P < 0.01$ for C.L.A. and the poorest-performing non-psychopathic inmate, respectively]. The psychopathic inmates performed like J.S., judging the conventional transgressions as very serious. In addition, unlike C.L.A. and the non-psychopathic subjects, J.S. judged that in the absence of prohibiting rules, both moral and conventional transgressions were not permissible. Chi-square analyses revealed that J.S. was significantly less likely to distinguish the moral and conventional transgressions in the absence of prohibiting rules ($\chi^2 = 13.8$ and 10.3, $P < 0.01$ for C.L.A. and the poorest-performing non-psychopathic inmate, respectively]. The psychopathic inmates, like J.S., judged both moral and conventional transgressions to be non-permissible in the absence of prohibiting rules.

Comment. J.S. showed profound impairment on the moral/conventional distinction test. J.S., like the comparison populations, judged the moral and conventional transgressions to be equally non-permissible. This suggests some access to 'social schema knowledge'; J.S. knows which behaviours society prohibits. However, unlike C.L.A. and the non-psychopathic inmates, J.S. judged the conventional transgressions to be as serious as the moral transgressions and considered the conventional transgressions to be non-permissible even in the absence of prohibiting rules. In these respects, J.S.'s and the psychopathic inmates' performance was similar.

Table 8 *Performance on the two social situations tasks (standard deviations and ranges in parentheses)*

	J.S.	C.L.A.	Healthy controls	Psychopathic inmates	Non-psychopathic inmates
Social Situations Task (Task 11)					
Normative situations identified (Max = 11)	9	8	9.3 (1.34, 7–11)	8.75 (2.63, 6–11)	9.25 (2.06, 8–11)
Violations of norms identified (Max = 11)	1*,†,‡,¶	11	9.90 (1.37, 7–11)	9.75 (0.96, 9–11)	8.50 (1.29, 8–11)
Appropriateness score [Violations] (Max = 33)	4*,†,‡,¶	23	19.6 (6.36, 11–31)	13.00 (4.24, 9–18)	14.50 (5.74,10–22)
Revised Social Situations Task (Task 12)					
Normative situations misidentified (Max = 16)	12	13	12.5 (0.97, 11–14)	13.00 (0.82, 12–14)	13.00 (13–13)
Violations of norms identified (Max = 20)	8*,†,‡,¶	15	17.6 (0.97, 16–19)	15.25 (0.96, 14–16)	16.5 (0.71, 16–17)
Appropriateness score [Violations] (Max = 60)	11*,†,‡,¶	36	38.10 (3.76, 33–45)	28.75 (5.74, 22–34)	33.75 (2.63, 31–36)

*J.S.'s performance is significantly poorer than that of the lowest scoring control; †J.S.'s performance is significantly poorer than that of the lowest scoring psychopathic inmate; ‡J.S.'s performance is significantly poorer than that of the lowest scoring non-psychopathic inmate; ¶J.S.'s performance is significantly poorer than C.L.A.'s. χ^2 analyses were used for between-subject comparisons of violations identification. One-way ANOVAs were used for the appropriateness scores.

Task 11: social situations task

In Task 10, J.S. demonstrated intact knowledge about social rules but failed to distinguish moral and conventional transgressions. Successful performance on the moral/conventional distinction task has been considered to be a consequence of an emotional reaction to moral transgressions. Moral transgressions result in victims. The sadness of these victims acts as an aversive stimulus. Representations of moral transgressions come to be associated with this sense of aversion through classical conditioning (Blair, 1995).

However, there are other emotional reactions that influence our judgements of the appropriateness of behaviour. There are acts for which there are no formal societal prohibitions but which may induce anger or irritation in observers; for example intimately touching another's child. In Task 11, we investigated whether J.S. was able to judge the appropriateness of behaviours that may induce anger in observers.

Procedure. Nine short stories describing social situations incorporating behaviour that was either normative or a violation were read by the patient (see Dewey, 1991). At various points in each story, the patient was asked to comment on how appropriate the behaviour was, giving a score from A to D. 'A' scores meant that he judged the situation as normative. 'B' to 'D' scores meant that he judged the situation as a norm violation and indexed the extent of the violation ('B' scores being mild and 'D' being serious). Previous piloting on a large, independent sample of healthy controls had resulted in the identification of a set of consistently identified normative situations and violations. This piloting had also indicated that the most commonly attributed emotion of observers to the violations was anger. The task instructions and an example story are presented in Appendix C (iv).

Three scores were obtained for this task (see Table 8).Two refer to the number of normative situations and the number of violations correctly identified. The third refers to the extent to which the patient judged the violations to be socially inappropriate. For each situation, the participant obtains a score between 0 and 3, matching their response of 'A' to 'D' (i.e. 'A' = 0, 'D' = 3).

Results. In Table 8, the performance of J.S. on the social situations task is contrasted with that of C.L.A., 10 healthy controls, and five psychopathic and five non-psychopathic inmates. J.S. was at the lower end of the healthy control range in his ability to identify normative situations. However, while C.L.A. and the healthy controls identified almost all of the norm violations, J.S. identified significantly fewer. His appropriateness score was significantly lower than those of the comparison populations.

Comment. J.S. showed no clear impairment in identifying normative situations. In contrast, he was strikingly impaired in identifying social violations. He was significantly poorer than the comparison populations in judging behaviours that are considered likely to induce anger in observers as being inappropriate. This was despite his intact knowledge of formal social rules (see Task 10).

Task 12: a revised social situations task

In the previous task, we obtained evidence that J.S. was impaired in judging the appropriateness of behaviours that may induce anger in observers. The purpose of Task 12 was to extend and replicate the results of Task 11 by using rather more extreme norm violations. These should engender even stronger expectations of anger in observers. Thus, the stories

in Task 12 dealt with issues such as violent aggression and public nudity [see Appendix C (v) for an example story].

Procedure. The procedure used in Task 12 was identical to that used in Task 11. Nineteen short stories describing social situations were read to the patient who was asked at various points in each story to comment on the appropriateness of the behaviour of the story protagonists.

Results. In Table 8, the performance of J.S. on the revised social situations task is contrasted with that of C.L.A., the healthy controls, and the psychopathic and the non-psychopathic inmates. J.S. again showed a preserved ability to identify normative situations correctly. However, unlike the comparison populations, J.S. was again seriously and significantly impaired in his ability to judge the actions likely to cause anger in others as being inappropriate.

Comment. The results of Tasks 11 and 12 indicated that J.S. was impaired in making behavioural appropriateness judgements reliant on the formation of expectations about the angry reactions of others.

Discussion

Following a bilateral trauma to the frontal region which involved orbitofrontal cortex, J.S. presented with severe behavioural disturbance and a profound impairment on a range of tests sensitive to frontal lobe dysfunction. His aberrant behaviour fulfilled the criteria for Antisocial Personality Disorder (DSM-IV, 1994) and can be described as 'acquired sociopathy' (Damasio, 1994); J.S. was notably aggressive and was reckless regarding the safety of others. We investigated the basis of his acquired sociopathy. In particular, we assessed the processes thought to be implicated in acquired sociopathy, i.e. reversal learning, expression recognition, emotional responding and social cognition. This allowed some insight into the specific impairments that may underlie acquired sociopathy and their interactions. Moreover, this study directly compared: (i) a dysexecutive patient with aberrant behaviour (J.S.) with a dysexecutive patient without aberrant behaviour (C.L.A.); and (ii) 'acquired sociopathy' with developmental psychopathy.

The results of the experimental investigation revealed that J.S. showed no impairment on two reversal learning tasks (Tasks 1 and 2). In contrast, both C.L.A. and the inmates with developmental psychopathy presented with impairments in one of the two tasks (Task 1 for C.L.A. and Task 2 for the inmates, respectively). J.S., despite a preserved performance on reversal learning tasks, showed a profound impairment on expression recognition and emotional responding tasks (Task 3 and 4). In contrast, patient C.L.A. presented with a selective deficit in the recognition of sadness. The inmates with developmental psychopathy presented with no emotion recognition impairment but a specific impairment in emotional responding to sad and fearful expressions. J.S.'s

profound expression recognition impairment could not be attributed to a general face processing impairment given his intact performance on a range of face processing tasks (Task 5). Nor could his emotional responding impairment be attributed to a general impairment in generating autonomic responses given his appropriate responding to auditory stimuli (Task 6).

In addition, J.S. showed impairment on most of the social cognition tasks. He failed to attribute fear, anger and embarrassment to others (Task 8). He failed to discriminate transgressions which result in victims (moral) and those which result in social disorder (conventional) (Task 10) and to judge inappropriate behaviours likely to induce anger in observers (Tasks 11 and 12). In stark contrast, his ability to attribute mental states on the Theory of Mind task was preserved (Task 9). Unlike J.S., C.L.A. showed no impairment on any of these tasks. Unlike J.S., the inmates with developmental psychopathy had no difficulty attributing emotional states to story protagonists and had no difficulty judging behaviours likely to induce anger in observers as being inappropriate (Tasks 11 and 12). However, like J.S., they also failed the moral/conventional distinction test (Task 10).

Alternative explanations for J.S.'s acquired sociopathy

J.S.'s behavioural disturbance and pattern of performance on the cognitive tasks cannot be explained easily within existing general accounts of aberrant behaviour/acquired sociopathy following frontal lobe damage. Damasio and colleagues have explained acquired sociopathy in terms of damage to the somatic marker system (e.g. Damasio, 1994). Damasio suggested that 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 causes behavioural disturbance. Two experimental findings have been used to support the somatic marker system model. The first is that patients who present with acquired sociopathy fail to show autonomic responses to visually presented social stimuli under passive viewing conditions (e.g. Tranel and Damasio, 1994). The second is that patients with acquired sociopathy perform poorly on the Four-Pack Card-Playing task (e.g. Bechara *et al.*, 1994). J.S.'s failure to generate autonomic responses to visually salient stimuli would be compatible with impairment in the somatic marker system. However, his preserved performance on the Four-Pack Card-Playing task would not. Moreover, J.S. showed reduced autonomic activity under *active* viewing conditions; when he was asked either to name the expression displayed or to award a score for the quality of the picture. These differences between J.S.'s pattern of performance and that reported for previous patients with acquired sociopathy indicate that J.S.'s impairment cannot be explained satisfactorily within the somatic marker system hypothesis.

Rolls related the aberrant behaviour to dysfunction in altering behaviour appropriately in response to reinforcement contingency changes (Rolls, 1996). However, J.S. did not appear impaired in the reversal learning tasks requiring the modification of behaviour according to reinforcement contingency changes. Thus, Rolls' account cannot easily interpret his behavioural disturbance.

Grafman and colleagues (e.g. Grafman, 1994; Grafman *et al.*, 1996) have argued that 'knowledge *stored* in the human prefrontal cortex plays a managerial role in the control of behavior and takes the form of mental models, thematic understanding, plans and social rules. . . . Within this framework, we would expect that lesions to the prefrontal cortex would impair the ability to access and sustain such managerial knowledge.' He has suggested that low rather than high frequency knowledge units would be more susceptible to degradation. In addition, he suggested that 'this impairment would bias the regulation and expression of behavior away from plans, social rules, and mental schemas towards environmental hyper-responsiveness, making spontaneously appearing or reactive aggressive and violent behavior more likely' (Grafman *et al.*, 1996, p. 1237). However, Grafman's position has difficulty accounting for J.S.'s pattern of performance. For example, J.S. showed difficulty with the moral/conventional distinction task. Yet all the transgressions described should be considered high frequency knowledge units as they concern common transgressions (e.g. one person hitting another, talking in class). Moreover, he clearly had access to the knowledge that the transgressions were prohibited. Rather, his impairment was in judging the conventional transgressions as *too* serious and not modulating his responding with respect to the conventional transgressions following the removal of the prohibitory rules.

Baron-Cohen has suggested that the impairment of patients with acquired sociopathy can be accounted for in terms of damage to the neural circuit that mediates Theory of Mind (Baron-Cohen, 1995). However, J.S. did not perform poorly on the Theory of Mind task. This suggests that the acquired sociopathy need not reflect Theory of Mind impairment. This is in line with the finding of Saver and Damasio that E.V.R. was unimpaired in a test that required representing potential mental states (Saver and Damasio, 1991).

Brothers implicated the orbitofrontal cortex in a neural circuit including the amygdala, anterior cingulate gyrus and temporal pole that functions as a unitary social 'editor' (Brothers, 1995, 1997). She argues against the idea of dissociating what she terms 'cold' social cognition (i.e. attributing beliefs; Theory of Mind) from 'hot' social cognition (i.e. processing others' emotional expressions). It is not clear to what extent Brothers would use the social editor model as a model of the acquired sociopathy shown by J.S. However, the data provided by J.S. have a clear implication for her position. J.S.'s pattern of performance indicates that the systems that are involved in 'cold' social cognition (i.e. Theory of Mind) do dissociate from those engaged in 'hot' social cognition (processing others'

emotional signals). Thus, it appears inappropriate to postulate a unitary social cognition module. Rather it appears that there are dissociable, perhaps interlocking, systems involved in social cognition.

In addition to the specific positions outlined above, there are at least two more general accounts that could offer explanations of J.S.'s data. The first is developed from the suggestion that the frontal lobes allow abstraction (e.g. Fletcher *et al.*, 1995). Indeed, J.S. (and C.L.A.) presented with concrete interpretations in the proverbs tasks. It could therefore be suggested that J.S.'s acquired sociopathy was due to an inability to abstract an appropriate behavioural plan from the environment. Thus, for example, J.S.'s impairment on the social situations tasks might be due to an inability to generate abstract representations of appropriate social behaviour. However, such an interpretation could not easily account for J.S.'s preserved performance on the Theory of Mind task which clearly requires the representation of an abstract mental state. In addition, C.L.A. also presented with an impairment in abstraction on the proverbs test, yet he had no difficulties on the social situations test.

A second general account could be offered in terms of an inhibition failure. The frontal lobes are thought to allow behavioural inhibition and regulation (e.g. Fuster, 1980; Shallice, 1988). Indeed, it has been suggested that violence in frontally impaired patients is due to the patients' inability to inhibit their violence impulses (e.g. Barratt, 1994; Krakowski *et al.*, 1997). However, a diminished inhibition account lacks specificity. We administered two tests thought to require the inhibition of a dominant response (Hayling and Stroop). J.S. showed impairment on the Hayling but intact performance on the Stroop task. Moreover, C.L.A., who showed no acquired sociopathy, was impaired on both tests. Of course, it would be possible to suggest that J.S. had suffered a deficit in a specific system that normally inhibited his aggression. However, this would be little more than a redescription of his symptomatology and would certainly not predict his pattern of impairment on the cognitive tasks.

A comparison between developmental and acquired psychopathy

Damasio has suggested that an impairment in the somatic marker system may be the underlying cause not only of 'acquired sociopathy' but also of developmental psychopathy (Damasio *et al.*, 1990; Damasio, 1994). This predicts that similarities should be observed between the two disorders. Indeed, the clinical and empirical picture of the developmental psychopath describes an individual who, like J.S., is highly aggressive and has an emotional deficit (reduced fear and guilt; e.g. Cleckley, 1964; Aniskiewicz, 1979; Hare, 1993; Patrick *et al.*, 1993; Blair, 1995).

Comparison of J.S.'s performance with that of individuals with developmental psychopathy revealed some striking differences in both form of aggression and type of emotional impairment. Individuals with developmental psychopathy are

known to show instrumental aggression. Even apparently motiveless attacks frequently are the result of imagined, or real, disrespect on the part of the victim (Hare, 1993). In contrast, J.S.'s aggression was never instrumental. It was usually a result of frustration at the interruption of his behavioural routines and, sometimes, for his entertainment (e.g. pushing the screaming woman around in her wheelchair).

J.S. and the psychopathic inmates also showed striking differences in their performance on the experimental tasks. J.S. had a severe problem in recognizing emotional expressions (particularly anger and disgust), showed severe and general deficits in autonomic responding and was impaired on all of the social cognition tasks with the exception of Theory of Mind. In contrast, the psychopathic inmates had no expression recognition deficit, showed a selective deficit in autonomic responses to sad and fearful expressions and did not show any deficit on the social cognition tasks, with the exception of the moral/conventional distinction task. These distinctive patterns cannot be attributed to severity of deficit, since the performance of the individuals with developmental psychopathy and that of J.S. overlapped depending on the experimental task. For example, the individuals with developmental psychopathy, but not J.S., were impaired on one task of reversal learning (the One-Pack Card-Playing task). Thus, these distinctive patterns of performance of J.S. and the individuals with developmental psychopathy do not support the notion that an impairment in the same cognitive system is the cause of both 'acquired sociopathy' and developmental psychopathy. It has been suggested that psychopathy is a potential developmental consequence of dysfunction within a neuro-cognitive system which responds to the sad and frightened faces of others (Blair, 1995; Blair *et al.*, 1997, 1999). Individuals with psychopathy do not find acts that cause harm to others aversive and fail to learn not to do these acts. The sudden onset of J.S.'s behavioural disturbance following his acquired brain damage makes the explanation of his aberrant behaviour in terms of a developmental dysfunction difficult. Instead, we believe that J.S. has suffered impairment to a system that is involved in the social on-line regulation of behaviour.

A new account of acquired sociopathy

So how can one explain J.S.'s performance? The essential features that marked J.S.'s impairment as distinct from the difficulties shown by C.L.A. and the individuals with developmental psychopathy were his impairment in recognizing, and responding to, angry expressions and his remarkably poor performance on the social situations tasks. Angry expressions are known to curtail the behaviour of others in situations where social rules or expectations have been violated (e.g. Averill, 1982). Thus, the suggestion here is that there is a system that is activated by another's angry expressions and which extinguishes on-going behaviour and reverses the current response in favour of another. It is suggested that this system may be activated by representations

of situations that previously have been associated with other individual's angry responses or other negative valence expressions (e.g. the staring expressions of others that can precede a sense of embarrassment and perhaps others' disgusted expressions), i.e. situations where another's anger might be expected. We suggest that J.S. may have suffered damage to such a system.

In terms of performance on the social situations tasks, we suggest that successful performance on these tasks requires that the representation of the situation activates an association with analogous situations where others have shown anger. For example, the Keith story [see Appendix C (ii)] has components of inappropriate intimacy and the touching of another's property that may initiate anger in others. Certainly, piloting indicated that most subjects anticipated that the baby's mother will be annoyed, or perhaps distressed, to discover an unknown man, Keith, fiddling with her baby's nappy. J.S., unlike healthy individuals, was unable to realize that some actions, such as checking a baby's nappy, while neutral or even reinforcing in many contexts (through the cessation of the baby's cries), are inappropriate in other contexts. We suggest either that J.S. did not activate the representations of situations that previously have been associated with other individual's angry responses or that these representations could no longer modify his on-going behaviour. Thus, J.S. did not realize the inappropriateness of the behaviours. This cannot be attributed easily to lack of knowledge about social rules or social plans. Indeed, there are no explicit social rules or social plans concerning these instances. Their inappropriateness is a function of the associations with other individual's reactions to the behaviour.

It should be noted that the above account may also be applicable to those patients who, following orbitofrontal cortex damage, presented with aggression (e.g. Grafman *et al.*, 1996). However, it is not intended to explain all the cognitive impairments that can be observed in patients with orbitofrontal cortex damage. For example, patient E.V.R.'s bankruptcy is more likely to have been due to a more general planning or decision-making deficit (Damasio, 1994). Similarly, the difficulties in shopping shown in two patients with orbitofrontal cortex lesions reported by Shallice and Burgess is likely to reflect dysfunction within a more general decision-making system (Shallice and Burgess, 1991). Indeed, it is important to note that neither E.V.R. nor the patients reported by Shallice and Burgess were documented as presenting with aggression or difficulties with anger management. We predict that it is only patients with these behaviours that will have impairment to the system described above. The orbitofrontal cortex is the third major division of the frontal cortex (Fuster, 1991). We suggest that the orbitofrontal cortex is involved not only in reversal learning to human social–emotional signals but also in a number of other cognitive processes.

Anatomical considerations

Although we are unable to delineate its exact extent, the CT scan revealed that J.S. had bilateral damage to the frontal

lobes involving the orbitofrontal cortex. Damage to the frontal lobes has long been implicated in causing a heightened risk of aggression (e.g. Blumer and Benson, 1975; Hecaen and Albert, 1978; Stuss *et al*., 1992; Damasio, 1995). Specifically, Damasio (e.g. Damasio, 1994), Grafman (Grafman *et al*., 1996) and Volavka (Volavka, 1995) have stressed the importance of orbitofrontal cortex in modulating appropriate social behaviour/inhibiting aggression. Also, Rolls and colleagues have reported generalized expression recognition impairment following orbitofrontal cortex damage (Hornak *et al*., 1996). In addition, a recent PET functional imaging study has suggested that the right orbitofrontal cortex is involved in the neural response to angry expressions (Blair *et al*., 1999). In line with this, J.S. showed generalized expression recognition difficulties. Moreover, when compared with the control subjects, his difficulties were notably striking for anger and disgust.

Furthermore, the neuropsychological evidence provided by the present case, in the context of previous findings, provides strong support for the functional specialization of the frontal cortex. Indeed, it suggests specialization of the frontal cortex for separable systems for social cognition. One system for social cognition allows the representation of the mental states of other individuals (Theory of Mind). Neuroimaging studies have indicated that the representation of the mental states of others incorporates medial frontal cortex (e.g. Fletcher *et al*., 1995; Goel *et al*., 1995; Gallagher *et al*., 2000). J.S. did not show any impairment in the representation of mental states. This suggests that his medial frontal cortex was intact. We propose a second system that responds to angry expressions/ expectations of other's anger. This system incorporates right orbitofrontal cortex (Blair *et al*., 1999). J.S. showed severe difficulty in processing angry expressions, which is consistent with his damage to this area.

J.S. also appears to have suffered damage to his left amygdala. The amygdala has been shown to respond to fearful (e.g. Adolphs *et al*., 1994; Calder *et al*., 1996; Morris *et al*., 1996; Phillips *et al*., 1997) and sad expressions (Anderson and Phelps, 1997, 1998; Blair *et al*., 1999). In line with this, J.S. showed impairment in processing both of these expressions. It has been suggested that the psychopathic disorder is a potential developmental consequence of early amygdala dysfunction (e.g. Blair *et al*., 1999; Blair and Frith, 2000). J.S. showed some overlap in cognitive impairment with the developmental psychopathic individuals. However, as argued above, this overlap in cognitive impairment was unlikely to underpin his socially aberrant behaviour. Yet it is interesting to consider whether the aggressive socially aberrant behaviour shown by J.S. and other cases might be dependent on damage to both orbitofrontal cortex and amygdala. In addition, it is interesting to consider whether there might be a form of developmental psychopathy that is due to orbitofrontal cortex damage and impairment in the system which responds to angry expressions/expectations of anger. These questions can only be addressed with further research.

Conclusions

As far as we are aware, this is the first study attempting to assess the processes implicated in the development of acquired sociopathy contemporaneously in the same patient. Moreover, it directly compared the performance of a patient with acquired sociopathy with cases of developmental psychopathy and with a patient with frontal dysexecutive syndrome. We are able to conclude the following. (i) Acquired sociopathy needs to be distinguished from developmental psychopathy. While both acquired sociopathy and developmental psychopathy may be associated with emotional impairments, we propose that the natures of these impairments are different. (ii) Acquired sociopathy need not be associated with general reversal learning impairments (see also Rolls, 1996). (iii) Acquired sociopathy is not an inevitable result of executive dysfunction even to 'inhibition' or 'abstraction' systems. We argue that the distinctive features of the 'acquired sociopathy' of J.S. were due to impairment to a system which responds to angry expressions/expectations of others' anger. This system may be particularly involved in the suppression of socially aberrant behaviour. Moreover, it appears that social cognition is not mediated by a unitary system. There appear to be multiple neurocognitive systems that are involved in social cognition. Given the importance of these systems for successful existence in society, their further delineation is clearly needed.

Acknowledgements

References

Adolphs R, Tranel D, Damasio H, Damasio AR. Fear and the human amygdala. J Neurosci 1995; 15: 5879–91.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. DSM-I. Washington (DC): American Psychiatric Association, 1952.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. DSM-III. 3rd edn. Washington (DC): American Psychiatric Association, 1980.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. DSM-III-R. 3rd rev edn. Washington (DC): American Psychiatric Association, 1987.

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. DSM-IV. 4th edn. Washington (DC): American Psychiatric Association, 1994.

Anderson AK, Phelps EA. Production of facial emotion following unilateral temporal lobectomy [abstract]. Soc Neurosci Abstr 1997; 23: 2113.

Anderson AK, Phelps EA. Bilateral amygdala damage impairs evaluation of facial but not vocal expressions of fear. Cognitive Neuroscience Meeting Abstract Program. 1998. p. 109.

Aniskiewicz AS. Autonomic components of vicarious conditioning and psychopathy. J Clin Psychol 1979; 35: 60–7.

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

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

Baron-Cohen S, Ring H, Moriarty J, Schmitz B, Costa D, Ell P. Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults. Br J Psychiatry 1994; 165: 640–9.

Barratt ES. Impulsiveness and aggression. In: Monahan J, Steadman H, editors. Violence and mental disorders: developments in risk assessment. Chicago: University of Chicago Press; 1994. p. 61–79.

Bechara A, Damasio AR, Damasio H, Anderson SW. Insensitivity to future consequences following damage to human prefrontal cortex. Cognition 1994; 50: 7–15.

Bechara A, Damasio H, Tranel D, Damasio AR. Deciding advantageously before knowing the advantageous strategy. Science 1997; 275: 1293–5.

Benton AL, Van Allen MW. Impairment in facial recognition in patients with cerebral disease. Cortex 1968; 4: 344–58.

Blair RJ. A cognitive developmental approach to morality: investigating the psychopath. Cognition 1995; 57: 1–29.

Blair RJ, Frith U. Neurocognitive explanations of the antisocial personality disorders. Criminal Behav Mental Health. In press 2000.

Blair RJ, Sellars C, Strickland I, Clark F, Williams AO, Smith M, et al. Emotion attributions in the psychopath. Personality Individual Differences 1995; 19: 431–7.

Blair RJ, Jones L, Clark F, Smith M. The psychopathic individual: a lack of responsiveness to distress cues? Psychophysiology 1997; 34: 192–8.

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

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: Grune and Stratton; 1975. p. 151–70.

Brothers L. Neurophysiology of the perception of intentions by primates. In: Gazziniga MS, editor. The cognitive neurosciences. Cambridge (MA): MIT Press; 1995. p. 1107–15.

Brothers L. Friday's footprint: how society shapes the human mind. New York: Oxford University Press; 1997.

Burgess PW, Shallice T. Response suppression, initiation and strategy use following frontal lobe lesions. Neuropsychologia 1996; 34: 263–72.

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

Calder AJ, Young AW, Rowland D, Perrett DI, Hodges JR, Etcoff NL. Facial emotion recognition after bilateral amygdala damage: differentially severe impairment of fear. Cogn Neuropsychol 1996; 13: 699–745.

Cleckley H The mask of sanity: an attempt to clarify some issues about the so-called psychopathic personality. 4th edn. St Louis (MO): Mosby; 1964.

Colby A, Kohlberg L. The measurement of moral judgement. Cambridge: Cambridge University Press; 1987.

Damasio AR. Descartes' error: emotion, reason, and the human brain. New York: Putnam; 1994.

Damasio AR, Tranel D, Damasio H. Individuals with sociopathic behavior caused by frontal damage fail to respond autonomically to social stimuli. Behav Brain Res 1990; 41: 81–94.

Damasio AR, Tranel D, Damasio HC. Somatic markers and the guidance of behavior: theory and preliminary testing. In: Levin HS, Eisenberg HM, Benton AL, editors. Frontal lobe function and dysfunction. New York: Oxford University Press; 1991. p. 217–29.

De Renzi E, Scotti G, Spinnler H. Perceptual and associative disorders of visual recognition. Neurology 1969; 19: 634–42.

Dewey M. Living with Asperger's syndrome. In: Frith U, editor. Autism and Asperger syndrome. Cambridge: 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.

Dimitrov M, Grafman J, Hollnagel C. The effects of frontal lobe damage on everyday problem solving. Cortex 1996; 32: 357–66.

Ekman P, Friesen WV. Pictures of facial affect. Palo Alto (CA): Consulting Psychologists Press; 1975.

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

Eslinger PJ, Damasio AR. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. Neurology 1985; 35: 1731–41.

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.

Fuster JM. The prefrontal cortex. New York: Raven Press; 1980.

Gallagher HL, Happé FGE, 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.

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

Grafman J. Alternative frameworks for the conceptualization of prefrontal lobe functions. In: Boller F, Grafman J, editors. Handbook of neuropsychology, Vol. 9. Amsterdam: Elsevier; 1994. p. 187–202.

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.

Happé FG. An advanced test of Theory of Mind: understanding of story characters' thoughts and feelings in able autistic, mentally handicapped, and normal children and adults. J Autism Dev Disord 1994; 24: 129–54.

Hare RD. The Hare Psychopathy Checklist—revised. Toronto: Multi-Health Systems; 1991.

Hare RD. Without conscience: the disturbing world of the psychopaths among us. New York: Simon and Schuster; 1993.

Hecaen H, Albert ML. Human neuropsychology. New York: 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.

Krakowski M, Czobor P, Carpenter MD, Libiger J, Kunz M, Papezova H, et al. Community violence and inpatient assaults: neurobiological deficits. J Neuropsychiatry Clin Neurosci 1997; 9: 549–55.

LeDoux JE. The emotional brain; the mysterious underpinnings of emotional life. London: Weidenfeld and Nicolson; 1998.

Leslie AM, Frith U. Autistic children's understanding of seeing, knowing and believing. Br J Dev Psychol 1988; 6: 315–24.

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

McKenna P, Warrington EK. Testing for nominal dysphasia. J Neurol Neurosurg Psychiatry 1980; 43: 781–8.

McNeil JE, Warrington EK. Prosopagnosia: a reclassification. Q J Exp Psychol 1991; 43: 267–87.

Morris JS, Frith CD, Perrett DI, Rowland D, Young AW, Calder AJ, et al. A differential neural response in the human amygdala to fearful and happy facial expressions. Nature 1996; 383: 812–5.

Nelson H. A modified card sorting test sensitive to frontal lobe defects. Cortex 1976; 12: 313–24.

Nelson HE, Willison J. The National Adult Reading Test. 2nd edn.Windsor (UK): NFER-Nelson; 1991.

Newman JP, Patterson CM, Kosson DS. Response perseveration in psychopaths. J Abnorm Psychol 1987; 96: 145–8.

Oldfield RC, Wingfield A. Response latencies in naming objects. Q J Exp Psychol 1965; 17: 273–81.

O'Sullivan M, Guilford J. Four factor tests of social intelligence (behavioural cognition): manual of instructions and interpretations. Palo Alto (CA): Consulting Psychologists Press; 1976.

Patrick CJ, Bradley MM, Lang PJ. Emotion in the criminal psychopath: startle reflex modulation. J Abnorm Psychol 1993; 102: 82–92.

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.

Phillips ML, Young AW, Senior C, Brammer M, Andrew C, Calder AJ, et al. A specific neural substrate for perceiving facial expressions of disgust. Nature 1997; 389: 495–8.

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

Raven JC. Advanced progressive matrices, sets I and II. London: H.K. Lewis; 1965.

Rolls ET. The orbitofrontal cortex. [Review]. Philos Trans R Soc Lond B Biol Sci 1996; 351: 1433–43.

Rolls ET, Hornak J, Wade D, McGrath J. Emotion-related learning in patients with social and emotional changes associated with frontal lobe damage. J Neurol Neurosurg Psychiatry 1994; 57: 1518–24.

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: 1241–49.

Shallice T. From neuropsychology to mental structure. Cambridge: Cambridge University Press; 1988.

Shallice T, Burgess PW. Deficits in strategy application following frontal lobe damage in man. Brain 1991; 114: 727–41.

Shallice T, Evans ME. The involvement of the frontal lobes in cognitive estimation. Cortex 1978; 14: 294–303.

Smetana JG, Braeges JL. The development of toddlers' moral and conventional judgements. Merrill-Palmer Q 1990; 36: 329–46.

Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935; 18: 643–62.

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

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

Tranel D, Damasio H. Neuroanatomical correlates of electrodermal skin conductance responses. Psychophysiology 1994; 31: 427–38.

Turiel E, Killen M, Helwig CC. Morality: its structure, functions, and vagaries. In: Kagan J, Lamb S, editors. The emergence of morality in young children. Chicago: University of Chicago Press; 1987. p. 155–243.

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

War Department, Adjutant General's Office. Army individual test battery: manual and directions for scoring. Washington (DC): War Department, Adjutant General's Office; 1944.

Warrington EK. Recognition memory test. Windsor (UK): NFER-Nelson; 1984.

Warrington EK The Camden Memory Tests. Hove (UK): Psychology Press; 1996.

Warrington EK, James M. Visual Object and Space Perception Battery. Bury St. Edmunds (UK): Thames Valley Test Company; 1991.

Warrington EK, McKenna P, Orpwood L. Single word comprehension: a concrete and abstract word synonym test. Neuropsychol Rehabil 1998; 8: 143–54.

Weigl E. On the psychology of so-called processes of abstraction. J Normal Social Psychol 1941; 36: 3–33.

Willison JR, Warrington EK. Cognitive retardation in a patient with preservation of psychomotor speed. Behav Neurol 1992; 5: 113–6.

Appendix A: case report of C.L.A.

C.L.A., a 49-year-old right-handed man, was admitted to the National Hospital in May 1998 for investigation of progressive motor difficulties which started ~2 years previously. On examination, he presented with left greater then right spastic quadraparesis, hyper-reflexia and extensor plantars. MRI revealed moderate supra and infratentorial atrophy. There were also extensive white matter changes in the right frontal lobe. The clinical diagnosis was motor neuron disease, probably primary lateral sclerosis.

Neuropsychological assessment

C.L.A. was tested on the shortened verbal version of the WAIS-R and obtained an average verbal IQ (see Table 1). On Raven's advanced progressive matrices, he obtained a good average score. Reading performance on the NART gave an estimated premorbid level of functioning in the high average range. On the Recognition Memory Test, he obtained a low average score on the verbal version and a defective score on the visual version. On the Topographical Memory Test, his performance was defective. There was no evidence of nominal difficulties. On the Graded Difficult Object Naming Test, he obtained a superior score. Visuoperceptual and visuospatial skills, as assessed by the Object Decision and Position Discrimination tests from the Visual Object and Space Perception Battery, were entirely within normal limits. On a series of speed and attention tasks, his performance was slow (Willison and Warrington, 1992).

Frontal executive functions

C.L.A. presented with no generalized behavioural difficulties. His personality and social interactions were judged as unchanged by his relatives since the onset of his illness. He never displayed aggression. However, there was some indication of emotional lability. He cried and laughed inappropriately on several occasions, indeed sometimes testing was interrupted by his inappropriate laughter. His performance was gravely impaired on a series of tests considered to be sensitive to frontal lobe damage. He was able to give only one of the two solutions on the Weigl Colour Form Sorting Test. On the Stroop Test, he was almost completely unable to name the colours of the printed words. He gave several bizarre responses on the Cognitive Estimates Test (e.g. how fast do race horses gallop = '. . . 65 miles per hour . . .'). On the WCST, he obtained only two categories (colour and shape), making several perseverative errors. On the Hayling Sentence Completion Test, he performed flawlessly on the response initiation section. However, on the response suppression section, almost all his answers were related to the sentence (all the guests had a very good . . .

'breakfast'). His verbal fluency was impaired. In 1 min, he produced only five words starting with the letter 'S' and seven food names.

Appendix B: population details of the psychopathic and non-psychopathic inmates

The five psychopathic and five non-psychopathic participants were all resident in Wormwood Scrubs prison. The participants were identified through their file records using the Revised Psychopathy Checklist (PCL-R; Hare, 1991). These files included reports by correctional officers, social workers and medical/psychiatric personnel, together with criminal records, court records, police and inmate versions of the offences, and regular reports on treatment progress/behaviour in prison and incidents involving the patient/prisoner. The PCL-R scores of the individuals were also calculated by independent clinicians resident at this institution. The inter-rater agreement with the PCL-R scores calculated by the first author was 0.95. There was 100% agreement in category (psychopathic or not) assignation. Subjects receiving a score of 30 or higher were classified as psychopathic (range 30–35), and subjects scoring 20 or below were classified as non-psychopathic (range 3–14). All of the subjects were serving life sentences for murder/manslaughter and had been in prison for at least 18 months. Life sentences for murder are mandatory. Life sentences for manslaughter are discretionary; they are given when there are no mitigating reasons for the act, such as a history of mental illness. None of the subjects were, or had been, psychotic or taking psychotropic medication. None of the subjects were comorbid with any Axis I (DSM-IV, 1994) disorder. All the subjects were male. The mean age of the psychopathic and non-psychopathic inmates was 48 years (SD = 13.02) and 49 years (SD = 11.90), respectively. The mean IQ of the groups was 91.2 (SD = 8.79) and 90.8 (SD = 5.89), respectively.

Appendix C: task details
(i) Task 8: examples of the stories used to elicit emotion attributions

Happy scene: Simon has just been told that his work within the company has meant that he will be winning an award.
Sad scene: Ol has been called in to his boss. His boss says 'You're fired.'
Fear scene: Gerald has driven into a safari park. His car has broken down and now the largest rhino is charging straight at him.
Anger scene: A man walks up to Larry and calls him an idiot.
Embarrassment scene: Ed is in a café when he slips up on some grease and falls straight over. Everyone in the café stares at him.

The participant is asked:
How will Simon/Ol/Gerald/Larry/Ed feel in this situation?

(ii) Task 9: example of a Theory of Mind story

Simon is a big liar. Simon's brother Jim knows this, he knows that Simon never tells the truth! Now yesterday Simon stole Jim's ping-pong bat, and Jim knows Simon has hidden it somewhere, though he can't find it. He's very cross. So he finds Simon and he says, 'Where is my ping-pong bat? You must have hidden it either in the cupboard or under your bed, because I've looked everywhere else. Where is it, in the cupboard or under your bed?' Simon tells him the bat is under the bed.

The participant is asked:

Q1: 'Was it true, what Simon told Jim?'
Q2: 'Where will Jim look for his ping-pong bat?'
Q3: 'Why will Jim look there for his bat?'

For the above story, an example mental states justification for the above story is 'Because Jim knows that Simon always lies and so he should look in the other location'. An example justification involving physical information for the above story is 'Because it will be in the opposite place to wherever Simon says'.

(iii) Task 10: moral/conventional distinction instructions and example scenes

Instructions: I am going to be describing to you some short scenes. All these scenes are set in a school. The scenes concern things that children frequently do at school. I am going to be asking you about your opinion of what is occurring in these scenes. There are no right or wrong answers. I am just interested in your opinion.

Example moral transgression scene:

One child runs up to another child and hits them in the face.

Example conventional transgression scene:

One child stands up and walks straight out of the classroom without permission in the middle of the lesson.

The participant is asked:

(1) 'Was it right or wrong for X to do Y?' (examining the participant's judgement of the permissibility of the act).

(2) 'On a scale of one to ten, how right (or wrong depending on the answer to 1) was it for X to do Y?' (examining the participant's judgement of the seriousness of the act).

(3) 'Why was it right (or wrong depending on the answer to 1) for X to do Y?' (examining the participant's theories about the act).

If the participant had said that the act was wrong in answer to question 1, the participant was then told: 'Now what if the teacher said before the lesson, before X did [the transgression], that 'At this school anybody can Y if they want to. Anybody can Y.' And then asked a final question: (4) 'Would it be O.K. for X to do Y if the teacher says X can?' (examining the rule's authority jurisdiction).

(iv) Task 11: social situations task instructions and example story.

Instructions: In the following stories some parts are in italics. Immediately following there is a pair of brackets ( ). Rate the behaviour which is illustrated by the portion in italics according to how you think most people would judge that behaviour if they witnessed it. Use this scale.

Fairly normal behaviour in that situation.	(A)
Rather strange behaviour in that situation.	(B)
Very eccentric behaviour in that situation.	(C)
Shocking behaviour in that situation.	(D)

Keith, age twenty-five, was a file clerk who worked in an office in the city. At noon he took his lunch to a small park and sat on a sunny bench to eat. *Often he tore part of a sandwich into bits, scattering it on the ground for pigeons.* ( ) One day when he came to his favourite bench a baby carriage was parked beside it. Keith noticed that a young woman was swinging an older child nearby. The baby in the carriage began to cry but the mother did not hear this because the swing was squeaking. Now, Keith had learnt that when his baby nephew screamed, sometimes this meant that a pin in his nappy had opened. *Rather than bother the mother in the park, Keith quickly checked the baby's clothing to see whether he could feel an open pin.* ( )

(v) Task 12: revised social situations example story

Robert is completely exhausted. He desperately wants to go to sleep. But he has to see his boss at his home that evening. He takes the tube to his bosses' house. *During the journey there, he rests his head against the glass window and falls asleep.* ( ) Half an hour later he wakes up and arrives at his bosses' station. Walking to the man's house makes him even more tired. His boss welcomes him and invites him in. Then his boss says that he will just be a few minutes, he has to sort something out upstairs. Robert goes into the main room of the house. There is a very thick carpet. He still feels so exhausted. *Robert lies down on the thick carpet and decides to have a nap.* ( )