Understanding social dysfunction in the behavioural variant of frontotemporal dementia: the role of emotion and sarcasm processing

C. M. Kipps,1,2 P. J. Nestor,2 J. Acosta-Cabronero,3 R. Arnold2 and J. R. Hodges2,4

1 Cognitive Disorders Group, Wessex Neurological Centre, Southampton University NHS Trust, Southampton, UK
2 Department of Clinical Neurosciences, University of Cambridge, Cambridge, UK
3 Wolfson Brain Imaging Unit, University of Cambridge, Cambridge, UK
4 Prince of Wales Medical Research Institute, Sydney, Australia

Correspondence to: Prof. J. R. Hodges
Prince of Wales Medical Research Institute
Barker St, Randwick
NSW 2031, Australia
E-mail: j.hodges@powmri.edu.au

Social interaction is profoundly affected in the behavioural form of frontotemporal dementia (bvFTD) yet there are few means of objectively assessing this. Diagnosis of bvFTD is based on informant report, however a number of individuals with a clinical profile consistent with the disease have no imaging abnormality and seem to remain stable, with doubt about the presence of underlying neurodegenerative pathology. We aimed to quantify aspects of the behavioural disorder and link it to the underlying level of atrophy in socially relevant brain regions. We tested individuals with either bvFTD (N = 26) or Alzheimer’s disease (N = 9) and 16 controls using The Awareness of Social Inference Test (TASIT) to assess their ability to identify emotion and sarcasm in video vignettes. A subset of bvFTD patients (N = 21) and controls (N = 12) were scanned using MRI within 6 months of assessment. There was marked impairment in the ability of bvFTD patients whose scans showed abnormalities to recognize sarcastic, but not sincere statements. Their capacity to interpret negative emotion was also impaired, and this appeared to be a major factor underlying the deficit in sarcasm recognition. Clinically diagnosed bvFTD patients whose scans were normal, Alzheimer’s disease patients and controls had no difficulty in appreciating both types of statement. In a multivariate imaging analysis it was shown that the sarcasm (and emotion recognition) deficit was dependent on a circuit involving the lateral orbitofrontal cortex, insula, amygdala and temporal pole, particularly on the right. Performance on a more global test of cognitive function, the Addenbrooke’s Cognitive Examination did not have a unique association with these regions. The TASIT is an objective test of social dysfunction in bvFTD which indexes the frontotemporal volume loss in bvFTD patients and provides an objective measure for separating behavioural patients who are likely to decline from those who may remain stable. These results provide additional evidence for the role of the orbitofrontal cortex and related structures in the processing of socially relevant signals, particularly those where negative emotion recognition is important.

Keywords: behavioural variant frontotemporal dementia; sarcasm; emotion; magnetic resonance imaging

Abbreviations: ACE = Addenbrooke’s Cognitive Examination; ADL = activities of daily living; bvFTD = behavioural form of frontotemporal dementia; BA = Brodmann Areas; CBI = Cambridge Behavioural Inventory; ES = effect size; FDR = false discovery rate; FTD = frontotemporal dementia; MNI = Montreal Neurological Institute; MMSE = Mini-Mental State Examination; NPI = Neuropsychiatric Inventory; TASIT = The Awareness of Social Inference Test; ToM = Theory of Mind; VBM = voxel-based morphometry
Introduction

Social dysfunction and abnormal behaviour are prominent features of frontotemporal dementia (FTD) syndromes. The behavioural form of FTD (bvFTD) is defined on the basis of impaired social interaction (Brun et al., 1994; Neary et al., 1998), but patients with semantic dementia (Bathgate et al., 2001; Snowden et al., 2001; Liu et al., 2004; Seeley et al., 2005; McMurtray et al., 2006; Mendez et al., 2006a), particularly where right temporal lobe atrophy predominates typically have similar problems (Miller et al., 1993; Edwards-Lee et al., 1997; Boone et al., 1999; Mychack et al., 2001; Perry et al., 2001; Thompson et al., 2003, 2004). A key objective in FTD research is to develop clinical instruments which are sensitive to the deficits in social cognition seen in FTD; yet are specific for the underlying neurodegenerative pathology and can be used for both early diagnosis and longitudinal disease tracking. A better understanding of the neural basis of the social deficit would also help in the management of these patients.

Diagnosis in bvFTD relies heavily on informant report and defined neuropsychological or imaging abnormalities are not mandatory features of the syndrome. A number of recent reports have highlighted the fact that a proportion of individuals fulfilling the accepted diagnostic criteria for bvFTD do not deteriorate and may not have a neurodegenerative syndrome (Davies et al., 2006a; Kipps et al., 2007c) but instead appear possibly to have either a developmental or neuropsychiatric disorder (see further discussion below). A handful of post-mortem studies have highlighted an absence of neuropathology in such individuals (Kertesz et al., 2005; Diehl-Schmid et al., 2007a), which is in keeping with preserved frontotemporal metabolism and structural integrity during life (Davies et al., 2006; Kipps et al., 2007c, 2008). The data on neuropsychological test performance in this group is limited, and no clear cognitive profile has, thus far been documented.

Tests of neuropsychological performance in bvFTD have tended to concentrate on aspects of executive function (Miller et al., 1991; Frisoni et al., 1995; Pachana et al., 1996; Kramer et al., 2003; Thompson et al., 2005): the ability to sustain and redirect attention, inhibit inappropriate responses and initiate or switch strategy (Alvarez and Emory, 2006). This reflects the perceived role of such tests as measures of frontal lobe function, which is regarded as the source of the disordered conduct. In bvFTD, however, both bedside cognitive assessments and standard formal neuropsychology may be normal (Gregory and Hodges, 1996; Gregory et al., 1998; Lough et al., 2001). One explanation is that these tasks may be insensitive to dysfunction of the medial and orbital frontal cortices, the right temporal lobe and amygdala which have been shown to degenerate early in bvFTD (Kril and Halliday, 2004; Kril et al., 2005).

Recently, tasks indexing emotional and more complex social dysfunction have been used in FTD, and several have been explicitly linked to an underlying neural substrate within the frontal and temporal lobes. Emotion processing is profoundly affected in FTD (Lavenu et al., 1999; Keane et al., 2002; Rosen et al., 2002, 2004; Fernandez-Duque and Black, 2005; Lavenu and Pasquier, 2005; Lough et al., 2006; Mendez et al., 2006b; Rosen et al., 2006), and typically involves a disproportionate failure in the ability to recognize negative emotions (anger, disgust, fear and sadness). This is associated with right hemisphere atrophy, particularly inferior and lateral aspects of the temporal lobe, right orbitofrontal cortex and the amygdala (Rosen et al., 2002, 2006). Most tests assessing emotion recognition have involved static photographs depicting facial emotion expression (Lavenu et al., 1999; Keane et al., 2002; Rosen et al., 2002, 2004; Fernandez-Duque and Black, 2005; Lavenu and Pasquier, 2005; Lough et al., 2006; Rosen et al., 2006; Diehl-Schmid et al., 2007b); one study using facial images which were blends of different emotions showed deficits in both positively and negatively valenced emotions (Kessels et al., 2007). The deficit is cross-modal, and abnormalities in the recognition of vocal emotion content (prosody) have also been seen (Keane et al., 2002).

Theory of Mind (ToM), the ability to ascribe beliefs, desires and intentions to another agent, is also abnormal in FTD as is self-referential processing and empathy, a wider concept involving the appreciation of others’ emotions (Lough et al., 2001; Gregory et al., 2002; Eslinger et al., 2005; Rankin et al., 2005; Lough et al., 2006; Rankin et al., 2006; Eslinger et al., 2007). Gregory et al. (2002) demonstrated variable deficits in bvFTD patients in first- and second-order ToM (e.g. X believes Y, or X believes that Z believes Y), however there were more marked impairments in the performance of ToM tests which involve emotion attribution, such as the faux pas test (Stone et al., 1998) (where the ability to sense another’s embarrassment is tested) and the Mind in the Eyes test (Baron-Cohen et al., 1997) (where mental state and emotion judgements are made on the basis of photographs of the eye region of models). The presence and extent of orbitofrontal atrophy, as assessed by a visual rating scale of coronal MRI scans, correlated with performance on the faux pas test. Subsequently, Torralva and co-workers (2007) showed that affective mental state attributions (as compared with cognitive) were most affected on the faux pas task. A marked deficit in empathic ability is present in FTD (Rankin et al., 2005; Lough et al., 2006), and may be related to the degree of atrophy of the right temporal pole, caudate nucleus and subcallosal gyrus in the medial prefrontal region (Rankin et al., 2006).

The precise role of executive function in social cognitive tasks remains controversial. Lough et al. (2006) showed that mental state attribution in a cartoon task in bvFTD patients was independent of the level of executive dysfunction, and suggested that this ability could be regarded as enabling, rather than defining, social cognitive performance. In contrast, Eslinger (2007) found that social judgements on a cartoon prediction task was determined by performance on the verbal-visual task (a simple measure of mental flexibility), second-order judgements on a standard ToM measure and empathy ratings, indicating perhaps a more important role for executive function. The disparity seen here in bvFTD patients echoes debate within the broader social cognition literature on the exact role of executive function in social cognition (Apperly et al., 2005). Possibly different social tasks involve executive functions to differing extents.
Although there is good evidence that the frontotemporal and amygdala regions degenerate in FTD (Broe et al., 2003; Kril and Halliday, 2004; Kril et al., 2005), most evidence to support a role for these structures in behaviour is based on subjective ratings of behaviour, such as apathy, disinhibition, changes in eating behaviour and stereotypic movements (Francesci et al., 2005; Rosen et al., 2005; Williams et al., 2005; McMurtry et al., 2006; Nakano et al., 2006; Peters et al., 2006). The data linking objective social cognitive deficits to these regions are, however, relatively sparse, and largely indirect, based on findings inferred from the functional imaging and lesion literature.

The orbitofrontal cortex is a richly interconnected region (Price 2006) with two functional networks: a medial aspect (incorporating the subgenual cingulate, claustrum and substantia innominata, and rostral amygdala), and a lateral aspect (BA 11, 12, 13) including the cingulate, paralimbic and limbic regions. In general terms, the orbitofrontal cortex implements rapid stimulus–reinforcer association learning which is particularly important for the dynamic changes implicit in reward processing and social interaction.

The orbitofrontal cortex, anterior insula and amygdala are strongly implicated in emotion and empathic processing on the basis of lesion studies (Adolphs et al., 1994; Rolls et al., 1994; Hornak et al., 1996; Calder et al., 2000; Adolphs, 2002; Hornak et al., 2003; Hornak et al., 2004; Shamay-Tsoory et al., 2004, 2005a, b). Hornak et al. (1996, 2003) showed that orbitofrontal lesions impair face and voice emotion identification (e.g. anger recognition (Murphy et al., 2003)), while bilateral lesions were also associated with more behavioural disturbance (particularly disinhibition) and a subjective change in emotional experience. Lesions of the ventromedial (orbitofrontal) cortex result in decreases in empathy (Shamay-Tsoory et al., 2005a) and poor performance on the faux pas task. A later study showed that affective rather than cognitive ToM is associated with the orbitofrontal region (Shamay-Tsoory et al., 2006). The medial prefrontal cortex appears to play a role in cognitive ToM judgements, and is activated by subjects thinking about self versus others (Kelley et al., 2002) and when considering how similar others are to oneself (Ochsner et al., 2005). It has been suggested that this region processes self-referential material from multiple domains (Amadio and Frith, 2006; Northoff et al., 2006). In the wider paralimbic belt, a wealth of data links the amygdala to fear processing (Adolphs et al., 1994; Adolphs, 2002), ToM ability (Stone et al., 2003) and to the capacity to perceive complex social emotions (e.g. guilt) (Adolphs et al., 2002); while multiple lines of evidence suggest a role for the anteroventral insula in the appreciation, and experience, of negative emotion such as disgust (Calder et al., 2000; Wicker et al., 2003; Kipps et al., 2007b).

Real-world interactions involve dynamic situations which are not fully captured by static images of emotion expression or morphed facial features, yet there are no studies in FTD which use dynamic social exchange as the basis of assessment. The Awareness of Social Inference Test (TASIT) uses trained method actors to portray exchanges depicting basic emotions, and more complex exchanges involving sarcasm, deception and irony (McDonald et al., 2003), and is well validated in controls and brain injured patients (McDonald et al., 2003, 2006).

The understanding of sarcasm is a complex process requiring appreciation of both the facts of a situation and the intention (mental state) of the speaker (McDonald, 1999). It is conveyed by the use of various paralinguistic emotional cues, such as facial affect and altered prosody, and is commonly intended to communicate criticism. Brain injured patients, with ventromedial (Shamay-Tsoory et al., 2005b) or orbitofrontal lesions (Channon et al., 2007), particularly in the right hemisphere perform poorly on tests of sarcasm detection which correlates with both empathic ability and the capacity for affective processing (facial expression and prosody) (Shamay-Tsoory et al. 2001, 2005b). This interaction of aspects of ToM and emotion processing in sarcasm detection is corroborated by a functional imaging study showing activation of the inferior frontal gyrus, temporal pole, superior temporal sulcus and medial prefrontal cortex, although all of the activations were in the left hemisphere (Uchiyama et al., 2006). In view of the regional atrophy of these structures in FTD, it would be expected that these patients would be poor at processing sarcastic stimuli.

We hypothesized that FTD patients with behavioural problems would have deficits in the processing of dynamic emotional interactions, and that this would in turn be correlated with impaired performance on a sarcasm task. Furthermore, we hypothesized that atrophy of the orbitofrontal cortex, particularly the lateral aspects involving BA 47/12, the temporal pole (BA 38) and the amygdala on the right would be particularly associated with this deficit, but that areas associated with executive functions not explicitly linked to social cognitive processing (dorsolateral prefrontal cortex) would show no association. In view of the previously mentioned reports of a clinically similar, yet possibly non-neurodegenerative copy of the bvFTD, we were particularly interested to see whether or not these individuals would show objective deficits on tests of social cognition, which would mirror their subjectively rated abnormalities of behaviour as this has not previously been reported. In view of their apparent lack of atrophy in frontotemporal regions, this might imply a more functional disturbance.

**Methods**

**Subjects**

Subjects (N = 51; FTD = 26, early Alzheimer’s disease = 9, controls = 16) were recruited from the Addenbrooke’s Early Onset Dementia and Mild Cognitive Impairment clinics. All patients are recruited prospectively and followed longitudinally with diagnosis based on accepted clinical criteria (Neary et al., 1998; Gregory et al., 1999; McKhann et al., 2001), and was not influenced by the results of social cognitive testing described below. The FTD patients all had gradual onset of behavioural and personality change as reported by a reliable informant (typically a spouse) with no psychiatric or other neurological explanation and the majority had been under review for a number of years. Early Alzheimer’s disease patients had prominent memory disturbance...
as the most salient feature with relative preservation of activities of daily living (ADL). Although their deficits were mild, all continued to decline in a manner typical of Alzheimer’s disease with longitudinal follow-up. Controls were age-matched to the clinical subjects. All subjects gave informed consent for the study according to the Declaration of Helsinki; in the case of patients, dual consent was obtained from a caregiver. The study was approved by the Local Research Ethics Committee.

Patients with bvFTD were divided into two groups on the basis of visual rating of raw structural images, performed blinded to any clinical data including test results and diagnosis according to our previously published rating scale which strongly predicts prognosis (Davies et al., 2006). Full details of the scale are available elsewhere (Kipps et al., 2007a), but briefly, this is a five-point scale assessing the degree of atrophy of the frontal lobes at the level of the temporal stem. A score of 0 or 1 overlaps the range seen in controls, and is regarded as normal, whereas a score of 2, 3 or 4 was never seen in any control and is regarded as abnormal. Of the 26 patients with FTD, 12 had defined MRI changes (designated FTDp—pathological), and 14 did not (designated FTDc—copy). It is important to reiterate that all patients with bvFTD had the clinical picture of the disorder with marked behavioural abnormality irrespective of the degree of atrophy noted on structural images.

**Behavioural stimuli**

The Emotion Evaluation and Test of Social Inference (Minimal) subtests from The Awareness of Social Inference Test (TASIT) (McDonald et al., 2007) were used to assess comprehension of basic emotion and the ability to detect speaker intention, attitude and meaning. The Emotion Evaluation subtest uses 28 professionally enacted video vignettes, with portrayals of positive (happiness, surprise and neutral) and negative emotions (anger, disgust, fear and sadness) lasting 20–30 s. Subjects were required to state the emotion portrayed by one of the actors in the vignette from a response card which included the emotions in random order. Each patient was queried as to their understanding of each emotion prior to the test commencing to exclude a relevant language deficit. All patients were easily able to identify gender from images of faces presented showing the eye region only, thus excluding a significant impairment of facial processing.

Fifteen video vignettes of actors making sincere, sarcastic or paradoxical sarcasm statements were then shown to subjects who were aware that they would subsequently be asked to endorse or reject a series of statements about what a specific actor was doing, saying, thinking and feeling. In the sincere exchanges, the targeted speakers mean what they are saying. In sarcastic exchanges, one of the speakers means the opposite of what he or she is saying and ‘intends’ the recipient to understand his or her real meaning. The dialogue for these scenes is identical to dialogue in ‘sincere’ scenes; therefore if the viewer is unable to detect sarcasm, they will misinterpret it as a sincere exchange. In clips with paradoxical sarcasm, the vignette does not make sense unless the viewer understands that one of the participants is being sarcastic. In these scenes, if the viewer does not detect the sarcasm, it is difficult for them to make sense of the exchange, and their answers are likely to be incorrect or bizarre [e.g. in one clip, one actor says (sarcastically) to the other that he has ‘torn up [his] ticket and thrown it away’ in response to a question about whether he has his ticket on him]. There was no restriction on the number of times subjects could watch any of the vignettes. If any question needed clarification, this was provided without explaining the emotion, sarcasm or mental states that were being portrayed. All vignettes were shown on a 17” computer screen, with an attached loudspeaker.

All subjects also had a Mini-Mental State Examination (MMSE) (Folstein et al., 1975) and the Addenbrooke’s Cognitive Examination (ACE) (Mathuranath et al., 2000; Mioshi et al. 2006). Functional status using the Clinical Dementia Rating (Morris, 1997) was recorded for patient groups (FTD, Alzheimer’s disease) and behavioural profiles were scored using the Cambridge Behavioural Inventory (CBI) (Bozeat et al., 2000) and the Neuropsychiatric Inventory (NPI) (Cummings et al., 1994).

**Analysis**

**Demographic data**

Demographic data were analysed using analysis of variance (ANOVA) (age, education, duration of illness at time of imaging, ACE and CBI) with post hoc contrasts. Non-parametric tests (Kruskal–Wallis and Mann–Whitney U) were used where appropriate (MMSE, CDR).

**Scores on the Emotion Evaluation Test**

Emotion recognition scores were combined to form two new variables: positive (happy and surprised) and negative (anger, disgust, fear and sadness) which were then normalized to a maximum score of 1. The data were not normally distributed and were analysed with non-parametric Kruskal–Wallis tests (positive versus negative emotion, group comparison) with post hoc testing corrected using Dunn’s Multiple Comparison Test.

**Scores on the Test of Social Inference (Minimal)**

Preliminary analysis suggested that scores for sarcasm and paradoxical sarcasm from the TASIT were similar, and so were combined to form a single sarcasm variable and analysed in a repeated measures ANOVA (within-subjects variable: statement type (sincere, sarcastic), between-subjects variable: group (FTDp, FTDc, Alzheimer’s disease and controls)).

**Imaging data**

**Participants and acquisition**

A subset of bvFTD patients ($N = 21$) and healthy controls ($N = 12$) had volumetric imaging within 6 months of their clinical testing. In this group, there were 11 bvFTD patients who had abnormal imaging (FTDp) on the semi-quantitative rating scale and 10 whose imaging was regarded as being normal (FTDc). Patients with Alzheimer’s disease were not included in this analysis.

A single 3D, spoiled gradient-recalled (SPGR) volumetric MRI was acquired for each subject. Imaging parameters were: $TE = 4.2 ms$, $TR = 13.5 ms$ with matrix size of 256 x 256 x 176 giving slice thickness of 0.98 x 0.98 x 1.5 cm.

**Preprocessing**

Images were preprocessed with N3 (McGill University, Montreal, Canada) and the Brain Extraction Tool (BET, FMRIB, Oxford, UK) as described in detail by Acosta-Cabronera et al (2008), and were visually inspected to ensure image quality before additional processing with SPM5 (Wellcome Dept Imaging Neuroscience, London, UK) which simultaneously normalizes and segments images into component tissue classes. Grey matter segments were modulated to sensitize the analysis to volumetric differences between scans; each voxel therefore reflects the concentration of grey matter at that location. An 8 mm full width at half maximum (FWHM) smoothing kernel was applied to the images.
Whole brain VBM analysis

An ANCOVA design using group membership as a factor (three levels) and global grey matter as a covariate was used to contrast volumetric differences between groups. Threshold masking was set at a relative threshold of 0.8 as a proportion of the global value. No regions of interest were specified for the volumetric contrasts. In view of the fact that the FTDp and FTDC groups had been divided on the basis of visual ratings, the purpose of this analysis was 2-fold: (i) to determine the regional profile of atrophy in the FTDp group relative to FTDC patients and controls, and (ii) to determine whether, despite their normal visual rating, whether FTDC patients had subtle atrophy relative to controls. The statistical threshold of significance was set at false discovery rate (FDR) <0.05 with a cluster threshold (K) >50 voxels.

Region of interest analysis

Based on previous work, our prior hypothesis was that grey matter density in six right hemisphere regions of interest were related to behavioural performance. The regions were (Fig. 1): R anterior cingulate (Montreal Neurological Institute (MNI) x,y,z centre of mass: 8,36,14), R medial orbitofrontal cortex (MNI 8, 50, −9), R lateral orbitofrontal cortex (MNI 41, 31, −13), R amygdala (MNI 27,−1, −19), R temporal pole (MNI 46, 13, −25) and L dorsolateral prefrontal prefrontal (MNI −34, 31, 34). These regions were chosen to reflect three separate postulated networks within the frontal lobes: lateral orbitofrontal-insula-amygdala, medial orbitofrontal-anterior cingulate-hippocampal and dorsolateral prefrontal-parietal association (Barbas and Zikopoulos, 2006; Price, 2006). In addition, two control areas were selected: the cerebellar vermis (MNI 2,−58, −19) as this is generally felt to atrophy relatively little in the disease (Kril et al., 2005), and the right temporo-parietal junction region (MNI 54, 24, 28) which is likely to atrophy less than frontotemporal regions, but may have a role in the representation of mental states (Saxe and Kanwisher, 2003). These coordinates were derived from the Automated Anatomical Labelling templates (Maldjian et al., 2003) within WFU Pickatlas (www.fmri.wfubmc.edu), a statistical parametric mapping (SPM) add-in module which is freely available.

Voxel values from these pre-specified regions of interest in modulated smoothed grey matter images, were extracted from the pre-processed SPM images using Marsbar (Brett et al., 2002) (http://marsbar.sourceforge.net) and entered into a multivariate analysis as dependent variables. Global grey matter volume was entered into the analysis as a nuisance covariate to adjust for the effects of brain size and global atrophy across the brain. In the first analysis, scores on the sincere and sarcastic conditions of the TASIT were simultaneously entered as additional covariates. A similar analysis was then performed using the composite score for negative emotion from the Emotion Evaluation subtest of the TASIT. In order to gauge the specificity of any observed differential effects of regions of interest (ROIs) in explaining the social cognition indices, a separate analysis assessed the effect of the ACE as a measure of global cognitive function.

In each case, main effects and interactions were followed up by univariate repeated measures ANOVAs contrasting the interaction by region against the effect on sarcasm of the cerebellar vermis. As a control region, the cerebellum is not thought to contribute to the behaviours under investigation, thus the magnitude of the effect size (ES) for each candidate ROI was ranked against this common denominator enabling ease of comparison of the contribution of each candidate region to the behavioural measures. Partial $\eta^2$ was used to report ESs for these analyses.

Results

Demographic data

There was no difference in the level of education or functional status (CDR), between patient groups, however, the FTD groups and controls were younger than Alzheimer’s disease patients $\{F(3,47)=2.90, P<0.05\}$, consistent with the typical later age of onset in these patients (Table 1). The FTDP group and Alzheimer’s disease patients were well matched on the MMSE (Mann–Whitney $U=51.5, Z=0.18, P>0.05$), but both groups performed worse than the FTDC group (FTDC versus Alzheimer’s disease: $U=10.5, Z=−3.4, P<0.01$; FTDC versus FTDP: $U=30.0, P<0.01$), who were themselves similar to controls.

Behavioural ratings using the CBI and the NPI were similar between patient groups for total scores, and CBI subscores for memory, ADLs or behaviour did not differ between groups, although the absolute levels of behavioural disturbance were much higher in the FTD groups. The profile of behavioural disturbance was qualitatively different in FTD and Alzheimer’s disease patients, but not in the two bvFTD subgroups.

Figure 1 Prespecified regions of interest were the right amygdala (purple, MNI coordinates x, y, z: 27, −1, −19), temporal pole (cyan, MNI 46, 13, −25), right lateral orbitofrontal cortex (mauve, MNI 41, 31, −13), right medial orbitofrontal and ventromedial cortex (light green, MNI 8, 50, −9), right anterior cingulate (red), left dorsolateral prefrontal cortex (yellow, MNI −34, 31, 34) and control regions cerebellar vermis (dark green, MNI 2, −58, −19) and right temporo-parietal junction (data not shown, MNI 54, 24, 28). Note: Shaded regions do not represent precise MNI co-ordinates used in ROIs.
Behavioural analysis

**Emotion evaluation**

**Negative versus positive emotions**

There was a group effect for negative [Kruskal–Wallis $H(3) = 21.37$, $P < 0.001$], but not positive emotions [Kruskal–Wallis $H(3) = 6.06$, $P = 0.05$]. Post hoc contrasts showed that the FTDp group performed worse than both the FTDc group (Wilcoxon $W = 16.33$, $P < 0.05$) and controls (W = 25.99, $P < 0.001$), but not Alzheimer’s disease patients. The FTDp group were also poor at recognizing positive emotions, and performed worse than FTDc ($W = 16.01$, $P < 0.05$) and Alzheimer’s disease patients ($W = 17.53$, $P < 0.05$), but not controls. FTDp patients were worse at recognizing negative emotions compared with positive ones ($W = 45$, $P < 0.01$), but for the other groups, performance across the two conditions was similar.

**Test of Social Inference**

A repeated measures ANCOVA with Greenhouse–Geisser correction was used for analysis of the TASIT (Mauchley’s W = 0.743, $\chi^2 = 13.675$, df = 2, $P < 0.01$). Since age and gender effect were possible, and there was a group difference in ACE scores, a series of exploratory two-way repeated measures ANCOVAs were performed with within-subjects factors: statement type (sincere, sarcastic) and, and group as a between-subjects factor (FTDp, FTDc, Alzheimer’s disease and controls) using age, gender and ACE scores as covariates. There was no main or interactive effect of these covariates in the model, either alone or in combination. The main analyses were therefore repeated without them. Significance thresholds are reported as $P < 0.05$.

There was a main effect of statement type [$F(1,47) = 14.89$, $P < 0.001$] which showed that sarcastic statements were harder to identify than sincerely expressed statements (Fig. 2). There was also a main effect of group [$F(3,47) = 10.74$, $P < 0.001$]. Post hoc contrasts showed that the FTDp group performed worse than all other groups ($P < 0.05$ versus FTDc, Alzheimer’s disease; $P < 0.01$ versus controls).

**Correlation between emotion recognition and sarcasm appreciation**

Performance on sarcasm recognition (SAR) correlated better with negative ($r = 0.70$, $P < 0.001$), versus positive (positive emotions...
emotion recognition (SAR: negative emotions $r = 0.53$, $P < 0.001$; positive $r = 0.23$, $P = 0.12$).

Table 2 MNI co-ordinates showing regions of atrophy in FTDp relative to FTDc and controls

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster size</th>
<th>FDR-corr ($P &lt; 0.05$)</th>
<th>$T$</th>
<th>$Z$</th>
<th>$x$</th>
<th>$y$</th>
<th>$z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>FTDp versus FTDc</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right caudate nucleus</td>
<td>4625</td>
<td>0.009</td>
<td>6.00</td>
<td>4.80</td>
<td>14</td>
<td>−12</td>
<td>20</td>
</tr>
<tr>
<td>Right temporal pole</td>
<td></td>
<td></td>
<td>5.90</td>
<td>4.75</td>
<td>22</td>
<td>10</td>
<td>−34</td>
</tr>
<tr>
<td>R inferior frontal gyrus</td>
<td></td>
<td></td>
<td>5.56</td>
<td>4.55</td>
<td>30</td>
<td>26</td>
<td>−24</td>
</tr>
<tr>
<td>Left temporal pole</td>
<td>2631</td>
<td>0.009</td>
<td>5.51</td>
<td>4.52</td>
<td>−48</td>
<td>18</td>
<td>−20</td>
</tr>
<tr>
<td>Left insula</td>
<td></td>
<td></td>
<td>5.42</td>
<td>4.47</td>
<td>−40</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>L temporal pole and Amygdala</td>
<td></td>
<td></td>
<td>5.11</td>
<td>4.28</td>
<td>−22</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>L middle temporal gyrus</td>
<td>54</td>
<td>0.012</td>
<td>4.21</td>
<td>3.69</td>
<td>−68</td>
<td>56</td>
<td>2</td>
</tr>
<tr>
<td>R fusiform gyrus</td>
<td>62</td>
<td>0.02</td>
<td>3.91</td>
<td>3.47</td>
<td>46</td>
<td>−44</td>
<td>−26</td>
</tr>
<tr>
<td>R inferior temporal gyrus</td>
<td>0.03</td>
<td></td>
<td>3.57</td>
<td>3.22</td>
<td>58</td>
<td>−48</td>
<td>−20</td>
</tr>
<tr>
<td>FTDp versus controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right insula</td>
<td>3514</td>
<td>0.007</td>
<td>6.18</td>
<td>4.90</td>
<td>40</td>
<td>0</td>
<td>−8</td>
</tr>
<tr>
<td>Right insula</td>
<td></td>
<td></td>
<td>6.18</td>
<td>4.90</td>
<td>34</td>
<td>20</td>
<td>−16</td>
</tr>
<tr>
<td>R inferior frontal gyrus</td>
<td></td>
<td></td>
<td>5.75</td>
<td>4.66</td>
<td>26</td>
<td>16</td>
<td>−20</td>
</tr>
<tr>
<td>Left temporal pole</td>
<td>1274</td>
<td>0.007</td>
<td>5.97</td>
<td>4.78</td>
<td>−36</td>
<td>16</td>
<td>−6</td>
</tr>
<tr>
<td>Left insula</td>
<td></td>
<td></td>
<td>5.62</td>
<td>4.59</td>
<td>−30</td>
<td>14</td>
<td>−16</td>
</tr>
<tr>
<td>Left temporal pole</td>
<td></td>
<td></td>
<td>5.06</td>
<td>4.25</td>
<td>−50</td>
<td>16</td>
<td>−14</td>
</tr>
<tr>
<td>Right precuneus</td>
<td>103</td>
<td>0.007</td>
<td>5.15</td>
<td>4.30</td>
<td>6</td>
<td>−54</td>
<td>68</td>
</tr>
<tr>
<td>R dorsolateral Prefrontal</td>
<td>113</td>
<td>0.008</td>
<td>4.87</td>
<td>4.13</td>
<td>60</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>R caudate nucleus</td>
<td>120</td>
<td>0.011</td>
<td>4.53</td>
<td>3.91</td>
<td>6</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>R caudate nucleus</td>
<td></td>
<td></td>
<td>3.71</td>
<td>3.33</td>
<td>10</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>R superior temporal gyrus</td>
<td>50</td>
<td>0.016</td>
<td>4.19</td>
<td>3.67</td>
<td>70</td>
<td>−42</td>
<td>20</td>
</tr>
<tr>
<td>Left fusiform region</td>
<td>50</td>
<td>0.029</td>
<td>3.70</td>
<td>3.32</td>
<td>−24</td>
<td>−8</td>
<td>−40</td>
</tr>
<tr>
<td>Left inferior temporal gyrus</td>
<td>3.48</td>
<td>3.15</td>
<td>−24</td>
<td>−4</td>
<td>2</td>
<td>−46</td>
<td></td>
</tr>
</tbody>
</table>

Local maxima and minima for SPM analysis comparing behavioural variant cases with and without atrophy as rated on semi-quantitative rating scale. Clusters are thresholded at FDR $P < 0.05$.

$T$ = t-statistic; $Z$ = z-statistic; $x,y,z$ = MNI coordinates of cluster.

Imaging results

Demographic data

In the 33 subjects (FTDp = 11, FTDc = 10, controls = 12) there was no difference in age [$F(2,32) = 0.99$, $P = 0.38$] or level of education [$F(2,28) = 0.92$, $P = 0.41$] at the time of MRI scanning, and the duration of illness was the same in the two FTD groups [$t(19) = 1.3$, $P = 0.18$].

Cognitive scores were different across the groups: on the MMSE [$H(2) = 10.27$, $P < 0.01$] the FTD group was worse than controls ($U = 18.0$, $Z = −3.0$, $P < 0.01$); FTDc patients and controls did not differ. There were significant group differences on the ACE [$F(2,31) = 2.85$, $P < 0.001$]; post hoc testing showed that the FTD group performed worse than the FTDc patients and controls, who were no different to each other.

Functional status (CDR) was similar in the two FTD patient groups, and on the CBI, endorsements of memory, ADL and behavioural subscales were identical. A similar result was obtained on the NPI.

Regional atrophy in FTDp, FTDc versus controls

There was marked volume loss (atrophy) in the FTD group relative to both controls and the FTDc group (see Table 2 for local maxima and cluster sizes and Fig. 3 for SPM showing distribution of atrophy), whereas no region was reduced in volume in the FTDc group relative to controls.

Emotion and sarcasm multivariate imaging analysis

Emotion evaluation

The multivariate analysis in which regional extracted voxel values (i.e. local volumes of grey matter) were tested as predictors of negative emotion scores revealed a strong effect of region [$F(7,22) = 3.31$, $P < 0.05$, ES = 0.5] and an effect of emotion [$F(1,28) = 7.99$, $P < 0.01$, ES = 0.22]. There was a region by volume interaction [$F(7,22) = 2.42$, $P = 0.05$, ES = 0.44], i.e. the volume of different regions differed across subjects, but there

\[ r = 0.3, P = 0.02 \]
was no region by emotion interaction \(F(7,22) = 1.42, P > 0.1, ES = 0.31\). However, in view of the ES of the interaction, post hoc contrasts were explored using unprotected univariate repeated measures ANOVA. There was a strong association between negative emotion recognition and the volume of the right amygdala \(P = 0.009, ES = 0.22\), right lateral orbitofrontal region \(P = 0.01, ES = 0.21\) and the right temporal pole \(P = 0.02, ES = 0.18\).

The right medial orbitofrontal cortex, left dorsolateral region and right temporo-parietal junction showed non-significant associations.

Sarcasm

When the combined sarcasm score from the TASIT was entered into a multivariate analysis with regional extracted voxel values as predictors there was a strong effect of region \(F(7,23) = 4.25, P < 0.004; ES = 0.56\) and of sarcasm performance \(F(1,29) = 18.34, P < 0.001, ES = 0.39\). Regional grey matter volumes also varied considerably across subjects \(F(1,29) = 109.87, P < 0.001\), and varied by region [for interaction: \(F(7,23) = 2.9, P = 0.03\)]. Importantly, there was a strong region \(X\) sarcasm performance interaction \(F(7,23) = 3.02, P = 0.02; ES = 0.48\). In other words, sarcasm recognition was strongly affected by the extent of regional brain atrophy. Follow-up contrasts for this interaction showed that, relative to the effect of cerebellar volumes on sarcasm performance, there was disproportionate atrophy of the right amygdala \(P = 0.003, ES = 0.26\), right lateral orbitofrontal cortex \(P = 0.01, ES = 0.19\) and right temporal pole \(P = 0.01, ES = 0.2\) in subjects who performed poorly on the sarcasm component of the test when controlling for their performance on interpretation of sincere statements. There was no region \(X\) sincere performance interaction effect \(F(7,23) = 0.98, P = 0.47\).

In summary, sarcasm performance was linked to the degree of atrophy in the right lateral orbitofrontal—temporal lobe—amygdala network, but not the medial orbitofrontal or dorsolateral prefrontal networks (Fig. 4).

Performance on the ACE showed a non-significant trend for an interaction with region \(F(7,23) = 2.02, P = 0.10\). Follow-up univariate analysis of the interaction contrasts, showed that relative to the cerebellum, all regions except the right temporo-parietal junction showed a relationship to performance (R amygdala, \(P = 0.001, ES = 0.31\); L-dorsolateral, \(P = 0.02, ES = 0.17\); R-anterior cingulate, \(P = 0.02, ES = 0.16\); R-lateral orbitofrontal, \(P = 0.001, ES = 0.31\); R-medial orbitofrontal, \(P = 0.01, ES = 0.21\); R-temporal pole, \(P = 0.002, ES = 0.29\); R-temporo-parietal junction, \(P = 0.17, ES = 0.07\)). In other words, ACE scores were affected by grey matter volumes in multiple brain regions including all three major sub-regions within the frontal lobes.

Post hoc extension of ROI

A post hoc extension of regions of interest to cover homologous regions in the contra-lateral hemisphere for each region was performed (left: amygdala, lateral orbitofrontal region, temporal pole, anterior cingulate, medial orbitofrontal cortex, and right: dorsolateral prefrontal cortex (for MNI coordinates, see Table 2)). In a multivariate analysis, there was a strong trend for an interaction between performance on sarcasm appreciation and regional volume \(F(13,17) = 2.15, P = 0.07\), but no interaction with sincere statement interpretation. Analysis of this interaction showed (ranked in order of ES): right amygdala \(F(1,29) = 10.42, P = 0.003, ES = 0.26\), right temporal pole \(F(1,29) = 7.34, P = 0.01, ES = 0.2\), right lateral orbitofrontal region/inferior frontal gyrus \(F(1,29) = 7.01, P = 0.01, ES = 0.19\), the left temporal pole \(F(1,29) = 6.79, P = 0.01\) and left.
The diagnosis of the bvFTD is a subjective one, made clinically without the need for imaging or neuropsychological abnormalities. Despite an identical behavioural symptom profile and duration of illness, FTD cases with imaging abnormalities (FTDp) were markedly worse on two social cognition tasks than patients who had normal imaging (FTDc) on a semi-quantitative rating scale, and patients with Alzheimer’s disease. Performance on these tests more accurately indexes brain atrophy, than individual symptoms in an unselected group of FTD patients, particularly within the orbitofrontal and temporal lobes. Using ecologically based tests of emotion and complex social behaviour, we were able to implicate anatomically relevant areas within the frontotemporal region in the social cognitive deficit observed in FTD patients.

It is notable that performance on the tests of social cognition was abnormal in the FTDp group only; that is to say, those patients who had imaging abnormalities despite an identical behavioural profile and no difference in the duration of illness compared with the FTDc group (those without imaging abnormality). If anything the FTDc patients had been symptomatic for longer. It has previously been shown that the FTDc group has a much better overall prognosis (Davies et al., 2006). The aetiology of FTDc is a matter of ongoing debate. It seems increasingly unlikely that they have an underlying neurodegenerative disorder given the lack of atrophy on MRI, the normality of metabolic brain imaging (FDG-PET) and the lack of progression to frank dementia even after many years of follow-up (Davies et al., 2006; Kipps et al., 2008). Yet they present with identical symptoms to those with FTDp as reported by reliable witnesses and have clearly undergone an insidious change in social behaviour and emotional regulation. A proportion probably fit within the Asperger spectrum and have decompensated in the mid to late life in the face of changing life circumstances. Others may have a neuropsychiatric syndrome with functional disruption of the same orbitofrontal-amygdala-polar network as undergoes degeneration in FTDp. Since their prognosis approximates to normal for their age there is currently a dearth of neuropathological data to confirm or refute these hypotheses.

Impaired performance on the sarcastic statements was not simply due to global cognitive decline in the FTDp group as performance on sincere statement recognition was normal in all groups. This finding has an important consequence: it suggests that the social cognitive deficit has no simple relationship with the typical symptom profile manifest by FTD patients as seen by the intact performance on social cognitive tasks by the FTDc group in the face of clear behavioural symptoms. Instead, it seems that abnormal social cognition depends on the degree of atrophy within the lateral orbitofrontal cortex and connected frontotemporal brain regions. It also suggests, however, that tests of emotion recognition and sarcasm may be useful in distinguishing FTD patients who are likely to progress rapidly from those with an FTD-like syndrome in whom there is unlikely to be neurodegeneration.

Sarcasm interpretation requires the integration of statement content with discordant paralinguistic cues, such as face expression or vocal prosody to accurately determine the speaker’s true intent. This is consistent with our finding that in general, sarcastic statements were harder to interpret than sincere statements which all groups performed without difficulty. Performance on sarcastic statements was strongly influenced by the ability to identify emotion, particularly negative emotion, from social interaction. This association remained even when taking into account more global measures of cognitive deficit, and suggests that the inability to recognize sarcasm stems from impaired emotion processing in bvFTD. As we did not specifically control for executive functions in this study, we do not discount their role in supporting the ability to block a literal interpretation of the sarcastic statements as to the speaker’s real intention, thus exposing the deficit in bvFTD patients’ performance.

A number of studies have reported impairments of sarcasm interpretation in patients with frontal brain regions.
(Shamay-Tsoory et al. 2001, 2005; Channon et al. 2007), and have suggested the importance of the orbitofrontal cortex as a neural substrate for this ability. None of these studies, however, have demonstrated a differential effect between the medial orbitofrontal (which is part of the medial network) and the lateral orbitofrontal cortex which has extensive interconnections with the insula, amygdala and temporal pole. Our study provides evidence to suggest that performance on sarcasm and emotion recognition tests is mediated by these regions, most notably in the right hemisphere. Disinhibition has been repeatedly shown as a consequence of orbitofrontal and inferior frontal lesions (Sarazin et al., 1998; Hornak et al., 2003), and the lateral orbitofrontal cortex is active in both emotional regulation and in social reversals (Hooker and Knight, 2006). We suggest that the integration of emotional content of sarcasm stimuli by the lateral orbitofrontal cortex is impoverished in bvFTD as a consequence of regional atrophy. Furthermore, impaired suppression of a dominant response (i.e. literal or non-sarcasitic interpretation) to the stimuli may additionally worsen performance. Damage to the amygdala and temporal poles would have the effect of further degrading the emotional valence and context of the sarcastic interaction (Olson et al., 2007). At least one functional imaging study has shown that the anterior cingulate is active in sarcasm interpretation (Uchiyama et al., 2006), however we did not demonstrate involvement of this region despite the need for mental state attribution in sarcasm interpretation, and the accepted role of the medial prefrontal cortex in mentalizing (Amodio and Frith, 2006). Perhaps this reflects the relatively mild atrophy within this region in our cohort, although at least one study has questioned the role of the medial prefrontal cortex in ToM ability (Bird et al., 2004).

Our study has demonstrated that it is not sufficient to simply rate behavioural abnormalities in bvFTD; while this has a clear role in clinical diagnosis, it does not identify subjects with underlying brain atrophy. It highlights the importance of reviewing imaging findings and employing a broader range of neuropsychological tests than has typically been the case in the assessment of these patients and assessment should include tests of social cognition, such as the TASIT with dynamic, socially relevant, stimuli. Future studies should attempt to determine whether these tests are sensitive at the earliest stages of disease, and whether they are able to adequately track progression of disease over time in individuals.

**Funding**

Wellcome Trust (073580, to C.M.K.); MRC-UK (to P.J.N.); Australian Research Council Federation Fellowship (FF 0776229, to J.R.H.).

**References**


Saxe R, Kanwisher N. People thinking about thinking people. The role of the temporoparietal junction in ‘‘theory of mind’’. Neuroimage 2003; 19: 1835–42.


