

Neuronal correlates of theory of mind and empathy: A functional magnetic resonance imaging study in a nonverbal task

Birgit A. Völlm,^{a,*} Alexander N.W. Taylor,^a Paul Richardson,^a Rhiannon Corcoran,^b
John Stirling,^c Shane McKie,^a John F.W. Deakin,^a and Rebecca Elliott^a

^aNeuroscience and Psychiatry Unit, University of Manchester, Stopford Building, Oxford Road, Manchester M13 9PT, UK

^bDepartment of Psychology, University of Manchester, Manchester M13 9PL, UK

^cDepartment of Psychology and Speech Pathology, Manchester Metropolitan University, Manchester M13 0JA, UK

Received 11 January 2005; revised 17 June 2005; accepted 8 July 2005

Available online 24 August 2005

Theory of Mind (ToM), the ability to attribute mental states to others, and empathy, the ability to infer emotional experiences, are important processes in social cognition. Brain imaging studies in healthy subjects have described a brain system involving medial prefrontal cortex, superior temporal sulcus and temporal pole in ToM processing. Studies investigating networks associated with empathic responding also suggest involvement of temporal and frontal lobe regions. In this fMRI study, we used a cartoon task derived from Sarfati et al. (1997) [Sarfati, Y., Hardy-Bayle, M.C., Besche, C., Widlocher, D. 1997. Attribution of intentions to others in people with schizophrenia: a non-verbal exploration with comic strips. *Schizophrenia Research* 25, 199–209] with both ToM and empathy stimuli in order to allow comparison of brain activations in these two processes. Results of 13 right-handed, healthy, male volunteers were included. Functional images were acquired using a 1.5 T Phillips Gyroscan. Our results confirmed that ToM and empathy stimuli are associated with overlapping but distinct neuronal networks. Common areas of activation included the medial prefrontal cortex, temporoparietal junction and temporal poles. Compared to the empathy condition, ToM stimuli revealed increased activations in lateral orbitofrontal cortex, middle frontal gyrus, cuneus and superior temporal gyrus. Empathy, on the other hand, was associated with enhanced activations of paracingulate, anterior and posterior cingulate and amygdala. We therefore suggest that ToM and empathy both rely on networks associated with making inferences about mental states of others. However, empathic responding requires the additional recruitment of networks involved in emotional processing. These results have implications for our understanding of disorders characterized by impairments of social cognition, such as autism and psychopathy.

© 2005 Elsevier Inc. All rights reserved.

Introduction

Human beings are profoundly social animals. The success of social interaction depends on the ability to detect cognitive and

emotional processes in others. The attribution of mental states, such as desires, intentions and beliefs, to others has been referred to as ‘theory of mind (ToM)’ or ‘mentalizing’ (Frith, 1999). Empathy, on the other hand, has been described as the ability to infer and share the emotional experiences of another (Gallese, 2003).

Several cognitive theories of ToM and empathy have been put forward. Premack and Woodruff (1978) suggested that the acquisition of theory of mind capacity is based on the development of a theory, e.g. a set of principles used to explain and predict phenomena. This account has become known as theory–theory and is advocated by Davies and Stone (1995), Gopnik and Astington (1988) and others. According to theory–theorists, experience provides infants with information that cannot be accounted for by their present theory of mind and this will eventually cause them to revise and improve that theory. According to the simulationist account (e.g. Gordon, 1986; Heal, 1995), the basis of theory of mind is an ability to imaginatively ‘put ourselves in another person’s shoes’, i.e. using our own minds to simulate the mental processes that are likely to be operating in the other. In this simulation theory, ToM is considered as innate and intuitive rather than the result of learning and experience. The centrality of emotion distinguishes empathy from other kinds of social cognition. Both the theory–theory and the simulation theory have been applied to empathy. According to Goldman (1995) empathy can be seen as a special case of simulation, where the output states are emotions. The theory–theory approach on the other hand suggests that we use knowledge, i.e. in the form of memories, in order to understand the situation of another. While the relative significance of these two theories remains uncertain, recent neurobiological research, in particular suggestions of a genetic locus for social cognition (Skuse et al., 1997) and the discovery of mirror neurons (Williams et al., 2001), might add weight to the simulation theory.

Observations in patient groups lacking ToM and/or empathy as well as recent neuroimaging research have provided empirical evidence for a neural basis of mentalizing and empathy. Impairment in theory of mind has been implicated in neurodevelopmental

* Corresponding author. Fax: +44 161 2757429.

E-mail address: birgit.vollm@man.ac.uk (B.A. Völlm).

Available online on ScienceDirect (www.sciencedirect.com).

disorders, particularly autism (Frith, 2001a,b), and in patients with certain circumscribed brain injuries (Stuss et al., 2001; Stone et al., 2003). Several brain imaging studies have investigated the neural basis of mentalizing in normal volunteers (Fletcher et al., 1995; Goel et al., 1995; Baron-Cohen et al., 1999; Brunet et al., 2000; Castelli et al., 2000; Gallagher et al., 2000, 2002; Russell et al., 2000; Vogeley et al., 2001; McCabe et al., 2001; Ferstl and von Cramon, 2002). Despite the use of different imaging techniques and contrasting cognitive activation paradigms, such as verbal and nonverbal tasks, these studies have produced remarkably consistent results. They reveal a network of three main areas associated with the processing of ToM stimuli including the medial prefrontal cortex (mPFC), the posterior superior temporal sulcus (STS) and the temporal poles (Frith and Frith, 2003). It has been proposed that the activation of these areas during mentalizing can be understood in terms of an evolution from preexisting processes such as the distinction between animate and inanimate entities, the self and the other and the ability to represent goal-directed actions (Gallagher and Frith, 2003).

Lack of empathy has been described in a group of individuals referred to as ‘psychopaths’ but little is known about the neural correlates of this deficit (Soderstrom, 2003). Few brain imaging studies in healthy volunteers have investigated the neuronal anatomy of empathy. Farrow et al. (2001), using fMRI, found activations in the superior and inferior frontal gyri, the precuneus and the middle temporal gyrus when respondents had to make empathic judgments in a verbal task. Using affective pictures evoking empathy, Moll et al.’s (2002) results indicated a role for the orbitofrontal, the medial frontal and the superior temporal gyri in the processing of moral emotions. Carr et al. (2003) described a network of activation including the superior temporal and inferior frontal cortices in an emotional observation task. Enhanced activations in these areas were found when subjects were asked to imitate the emotions displayed. Thus, similar brain areas seem to be activated when people infer and when they share emotions. Singer et al. (2004) compared brain activity while healthy volunteers experienced pain to that when they observed a loved one receiving a painful stimulus. Anterior insula and anterior cingulate were activated in both conditions and correlated with empathy scores while activations in the somatosensory cortex were only seen when participants received pain.

The relationship between ToM and empathy has yet to be determined. While the simulation theory supposes that simulation supports both processes, others have argued for a dissociation of ToM and empathy (Blair et al., 1996). The latter theory is based on the evidence of impairment of social cognition in autistic and psychopathic populations: while autistic individuals profoundly lack ToM skills, they show physiological and to some extent cognitive responses to distress cues in others. Psychopaths on the other hand are incapable of empathizing while there is no impairment of mentalizing.

Based on the available evidence from brain imaging studies, ToM and empathy seem to engage common as well as distinct neuronal networks. However, no studies have been carried out linking these two processes. Such an approach could help to further our understanding of associations between ToM and empathy. In the present fMRI study, our aim was to investigate the neuronal correlates of ToM and empathy in healthy volunteers using a visual cartoon task. In this context, we refer to empathy as the attribution of emotion to another individual. Subjects had to make inferences about either the mental or emotional state of the story protagonists.

We hypothesized that both conditions would activate temporal lobes and medial prefrontal cortex, areas implicated in social cognition (Adolphs, 2003) as well as perspective taking (Ruby and Decety, 2004). We further hypothesized that processing of empathic stimuli would also rely on affective networks, particularly the amygdala.

Methods

Subjects

Fifteen right-handed, healthy male participants were recruited from the general and university student population. Two participants were later excluded from the study due to high error rates in the task used, i.e. their scores were outside a margin of ± 2 standard deviations from the average error rate. Hence, results are reported for 13 subjects. Psychiatric morbidity and substance use disorders were screened for using the Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1997). None of the participants had any current or past serious medical or neurological condition or had taken any medication or illicit substances in the 2 weeks prior to scanning. The mean age was 24.9 years (range = 19–36 years). IQ was within the normal range with a mean of 106 (range = 94–122) as tested using the Quick Test (Ammons and Ammons, 1962). The study was approved by the local Ethics Committee and all participants gave written informed consent.

Activation paradigm

A visual activation paradigm comprising a series of comic strips, each depicting a short story, was used. Stimuli were presented in blocks. There were four categories of stories:

1. ‘Theory of Mind’ (‘ToM’);
2. ‘Empathy’ (‘Emp’);
3. ‘Physical causality one character’ (‘Physical 1’) and
4. ‘Physical causality two characters’ (‘Physical 2’).

The fMRI task was derived from Sarfati et al. (1997) and Brunet et al. (2000). Brunet et al., in their PET study used three experimental conditions: ‘Attribution of intention’ (i.e. ToM), ‘Physical causality with characters’ and ‘Physical causality involving only objects’. The authors showed that the ToM component of this task was associated with a pattern of brain activity involving medial and inferior prefrontal cortex and temporal lobes. For our ‘ToM’ blocks, we used the original cartoons from Brunet et al.’s (2000) ‘Attribution of intention’ condition. Some of the stimuli of the ‘Physical causality one character’ condition were also derived from Brunet et al.’s (2000) original cartoons. The remainder were drawn by an artist closely matching the style and complexity of these original cartoons. The ‘Physical causality two characters’ and the ‘Empathy’ stimuli were developed in our department.

‘Empathy’ stimuli were piloted in a group of 15 subjects outside the scanner. Subjects had to rate each cartoon for clarity and empathic understanding on a scale from 1–5 (very poor, poor, average, good and excellent). The pilot task was introduced by the following instruction: “The cartoons that will be presented require you to put yourself in the situation of the main character.” Specific questions asked were: “How good was your understanding of the cartoon strip in the time allowed?” for clarity and “How clear was

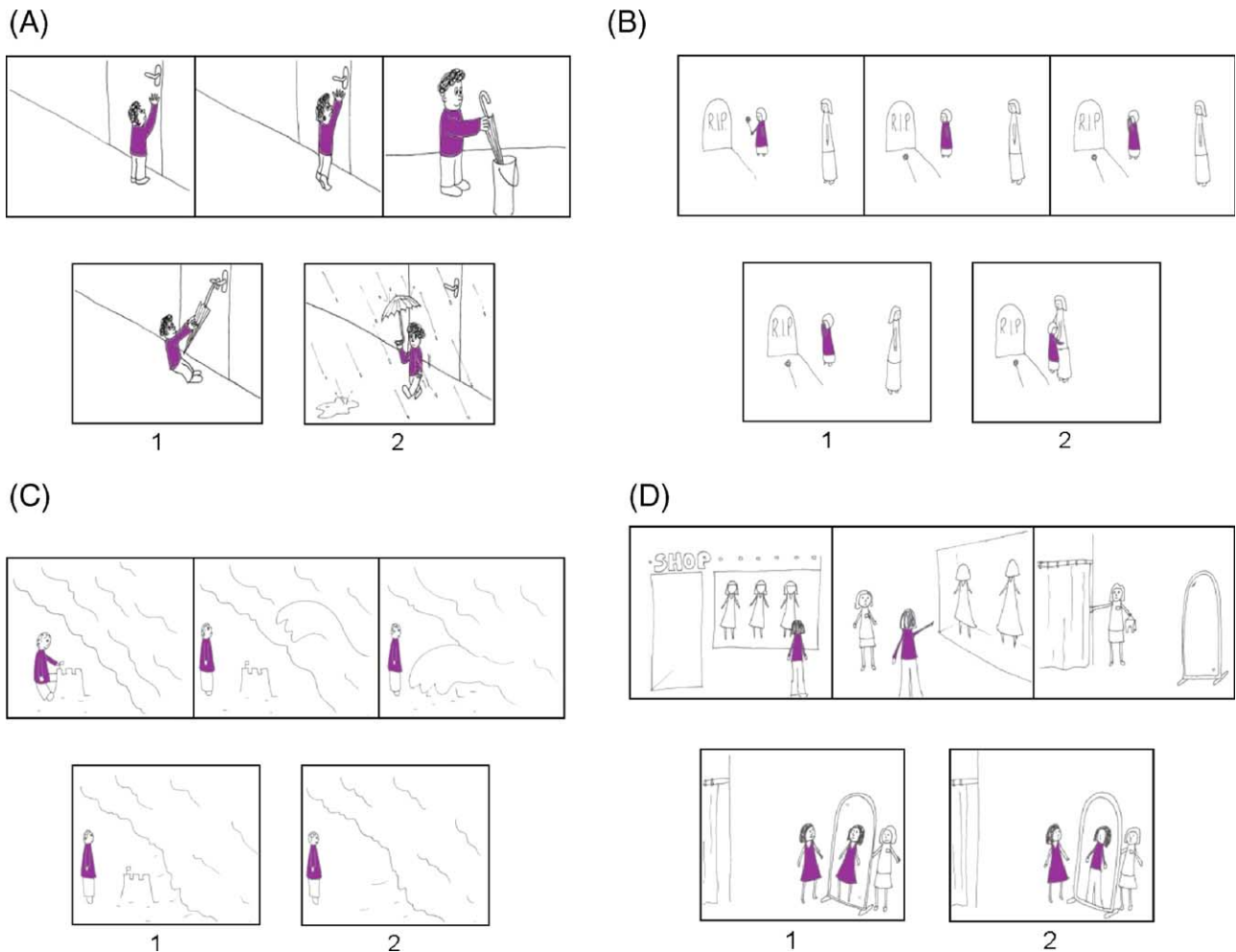


Fig. 1. Example of a story stimuli. (A) Theory of mind story. Picture 1 represents the 'correct' answer. (B) Empathy story. Picture 2 represents the 'correct' answer. (C) Physical causality one character story. Picture 2 represents the 'correct' answer. (D) Physical causality two characters story. Picture 1 represents the 'correct' answer.

your feeling for the main character's emotions?" for empathic understanding. Participants were also asked to summarize the plot and describe the feelings of the main character. Cartoons that seemed to lack clarity on the basis of this pilot were redrawn. Average rating scores of the finally included stories were 4.1 for clarity and 3.9 for empathic understanding indicating that the cartoons were easy to comprehend and engaged empathy for the main story character. Participants provided sensible descriptions of the story line and used emotional terms to describe feelings of the main character.

Each condition was presented twice so that the task consisted of eight blocks in total. In each block, five different comic strips depicting a short story were presented. Upon the second presentation of a condition, a new set of cartoons was used, hence each cartoon was only seen once. Each comic strip was shown for 6 s on the upper half of the screen. Then two pictures showing possible outcomes of the scenario were superimposed on the bottom half of the screen for a further 4.5 s. Participants were required to make a choice between these two story endings using a button box.¹ Only one of the outcomes represented a plausible

story ending ('correct ending'). At the beginning of each block, a short question introducing the block was shown for 6 s. The total block length was therefore 58.5 s, hence the task lasted 7 min and 48 s in total.

The introductory questions given at the beginning of each block were designed to engage the corresponding mental construct in the participant. In the ToM condition, the question was: "What will the main character do next?" The comic strips in this condition involved one character whose intentions had to be inferred by the subject in order to give the 'correct' story ending. In this condition, no social interactions or emotional situations were depicted. The empathy condition was introduced with the question "What will make the main character feel better?" Scenarios involved interaction between story characters and the correct answer required the volunteer to empathize with the protagonist. Both, the 'Physical 1' and 'Physical 2' conditions, relied on the comprehension of physical causalities only and were introduced with the prompt: "What is most likely to happen next?" In order to match the control and active conditions for number of characters and complexity, 'ToM' and 'Physical causality one character' stories depicted one character only whereas 'Physical causality two characters' and 'Empathy' stimuli depicted two characters. Examples of stimuli from each condition are shown in Fig. 1.

¹ In Brunet et al.'s (2000) activation paradigm, a choice of three possible endings was used.

Scanning procedure

Stimuli were rear-projected from a laptop computer onto a screen, which was visible via a mirror within the scanner. The screen was adjusted for each individual to optimize the visual field. Responses were given using a four-button button box of which only two buttons were required to choose the appropriate story ending. Responses were recorded.

Data acquisition

Functional magnetic resonance (fMRI) images were acquired using a 1.5 T Philips Gyroscan ACS NT (Philips, Best, NL) scanner. 95 volumes were acquired, each comprising 40 T2*-weighted contiguous axial slices (slice thickness = 3.5 mm), acquired using a single shot echo planar (EPI) pulse sequence (TR = 5 s; TE = 40 ms; in-plane resolution 3 mm × 3 mm).

Data analysis

Sociodemographic, psychometric assessment and test performance data were analyzed using SPSS 11.5. A General Linear Model, repeated measures was used to compare means of task performance data. All tests were two-tailed and a value of $P \leq 0.05$ was used to determine statistical significance.

fMRI data were processed using Statistical Parametric Mapping (SPM2) with a random effects model (<http://www.fil.ion.ucl.ac.uk/spm>). Individual scans were realigned using the first scan as a reference and normalized into the Talairach and Tournoux stereotactic space (Talairach and Tournoux, 1998) using the MNI templates. Spatial smoothing was applied with a 10 mm Gaussian kernel. Image data were high-pass filtered with a cut-off of 354 s.

Statistical analysis was carried out using the general linear model with a delayed boxcar waveform to model blood oxygenation level dependent (BOLD) signal changes to the 'ToM' or 'Empathy' condition, respectively, relative to the 'Physical 1' or 'Physical 2' conditions. In each condition, the full block length was modeled, including presentation and selection phase. Further, analysis of the conjunction of, and interaction between, the contrasts 'ToM' – 'Physical 1' and 'Empathy' – 'Physical 2' were conducted. Goodness-of-fit (beta) values for each contrast resulted in a contrast map for each individual. The statistical parametric maps from each individual data set were then entered into second-level, random effects analyses accounting for both scan-to-scan and subject-to-subject variability. The group effects of interest were assessed through applying one-sample t tests to the contrast map voxels. A voxel was deemed significant if its z score was greater than 3.09, corresponding to $P < 0.001$ uncorrected. While this is less rigorous than reporting only those voxels where BOLD response survived Bonferroni correction for multiple comparisons, the exploratory threshold of $P < 0.001$ uncorrected has been widely used in previous fMRI studies of subtle cognitive processes, particularly if there is a clear anatomical hypothesis as in our study. A cluster size of more than 10 contiguous voxels was applied.

Results

Task performance

Responses were recorded for all trials and failure to respond during the 4.5 s response window counted as an absent response.

As shown in Table 1, most stories were answered within the allowed response times. There were no significant differences in the mean number of answers given between conditions. Percentages of 'correct' answers of all given answers are shown in Table 1. An ANOVA, repeated measures revealed a significant effect of condition ('ToM', 'Empathy', 'Physical 1', 'Physical 2') on the percentage of correct responses ($P = 0.04$). Post-hoc paired samples t tests revealed a significant difference between 'Physical 2' and both 'Empathy' ($P = 0.001$) and 'Physical 1' ($P = 0.005$).

Functional imaging

Activations associated with the processing of theory of mind stimuli: 'ToM' – 'Physical 1'

In this comparison, neural responses in the 'Physical 1' blocks were subtracted from those in the 'ToM' blocks to reveal areas of increased signal associated with ToM processing.

Significant activations were found mainly in temporal and frontal cortices (Table 2). Temporal activations included bilateral temporoparietal junction, superior temporal sulci and temporal poles. Frontal areas of activation comprised bilateral orbitofrontal cortices. Further areas associated with the processing of 'ToM' stimuli included left fusiform gyrus, cuneus, posterior cingulate and bilateral cerebellum.

Activations associated with the processing of empathy stimuli: 'Empathy' – 'Physical 2'

In this comparison, neural responses in the 'Physical 2' blocks were subtracted from those in the 'Empathy' blocks to reveal areas of increased signal associated with the processing of empathy provoking stimuli.

This contrast was associated with statistically significant activations in the left temporoparietal junction and middle temporal gyrus/temporal poles bilaterally (Table 2). Activation clusters were also found in the left medial prefrontal cortex, left lingual gyrus extending to posterior cingulate and in the cerebellum bilaterally.

Common activations associated with the processing of 'Theory of mind' and 'Empathy' stimuli: conjunction analysis

We used a conjunction analysis to explore areas commonly activated by both the 'ToM' and 'Empathy' conditions (Table 2, Fig. 2). These areas included bilateral temporoparietal junction, middle temporal gyri extending to temporal poles and inferior temporal gyri extending to fusiform gyri. Frontal areas of activation were found in medial prefrontal and right orbitofrontal cortex. Further foci of activation associated with theory of mind and empathy processing were located in the middle and inferior occipital gyri including the left lingual and fusiform gyrus, and in the cerebellum.

Table 1
Task performance of 13 subjects in the behavioral ToM/empathy task

Condition	Mean number of answers given (maximum 10)	Correct answers (in % of all given answers)
'ToM'	9.3 (SD 0.9)	94.4 (SD 6.7)
'Empathy'	9.0 (SD 0.9)	98.4 (SD 4.0)
'Physical 1'	9.0 (SD 1.3)	96.6 (SD 6.8)
'Physical 2'	8.9 (SD 1.1)	84.1 (SD 1.3)

Table 2
Foci of significant brain activations associated with ToM and empathy stimuli

Area of activation	Brodmann area	<i>Left brain</i>					<i>Right brain</i>				
		z score	Cluster size	Tailarach coordinates			z score	Cluster size	Tailarach coordinates		
				x	y	z			x	y	z
<i>‘ToM’ – ‘Physical 1’</i>											
Orbitofrontal cortex	47	3.17	36	−48	28	−17	3.40	100	18	49	−8
Temporoparietal jct., extending to occipital gyrus	19/22/39	4.01	397	−59	−57	22	3.98	201	45	−78	18
Superior temporal gyrus	22						3.59	34	62	6	−5
Middle temporal gyrus, extending to temporal pole	21	3.35	119	−59	−21	−7					
Inferior temporal gyrus	21	4.39	71	−50	−7	−27	3.27	128	62	−15	−17
Posterior cingulate gyrus	31	3.39	217	0	−51	25					
Fusiform gyrus	37	3.16	36	−36	−45	−18					
Cuneus	18	4.04	45	−15	−102	0					
Cerebellum		3.15	23	−15	−89	−26	3.18	19	39	−72	−32
<i>‘Empathy’ – ‘Physical 2’</i>											
Medial prefrontal cortex	8	3.18	55	−9	37	51					
Temporoparietal junction	22/39	3.83	70	−50	−57	17					
Middle temporal gyrus/Temporal pole							3.17	36	48	4	−33
	21/38	3.11	56	−45	10	−31	3.36	56	57	3	−21
Middle occipital gyrus	18						3.87	11	27	−96	8
Lingualis gyrus, extending to posterior cingulate gyrus	19	4.47	129	−12	−52	3					
Cerebellum		4.13	19	−30	−80	−26	3.57	41	18	−83	−26
<i>Conjunction analysis</i>											
Medial prefrontal gyrus	8	3.48	155	−21	35	51	3.34	39	15	37	51
Medial orbitofrontal gyrus	25						3.59	31	6	8	−16
Temporoparietal junction	22/39	5.43	308	−53	−57	19	3.94	189	53	−60	17
Middle temporal gyrus, extending to temporal pole	21/38	4.08	234	−53	−9	−18					
Inferior/middle temporal, extending to fusiform gyrus	20/21	4.70	285	−50	−4	−30	4.33	304	59	−12	−15
Middle occipital gyrus/Cuneus	18/19	4.77	144	−27	−96	10	3.77	17	18	−101	11
Inferior occipital gyrus, extending to fusiform gyrus	18/19	4.09	146	−21	−102	−2	3.39	13	42	−87	4
Lingual gyrus	18	5.03	439	−9	−52	3					
Cerebellum		4.33	61	−30	−80	−26	4.41	89	21	−83	−26
<i>(‘ToM’ – ‘Physical 1’) – (‘Empathy’ – ‘Physical 2’)</i>											
Middle frontal gyrus	10	3.67	40	−39	55	3					
Orbitofrontal gyrus	11						3.65	220	48	46	−7
Superior temporal gyrus	22						4.02	57	62	3	−5
Occipital lobe/Cuneus	19						3.27	157	3	−89	29
<i>(‘Empathy’ – ‘Physical 2’) – (‘ToM’ – ‘Physical 1’)</i>											
Medial prefrontal cortex	10						3.39	13	9	53	14
Parahippocampal gyrus/Amygdala											
	28	3.28	27	−21	−1	−15					
Anterior cingulate	24/32						3.15	33	24	13	24
Posterior cingulate							3.12	15	6	−22	34
Middle occipital gyrus	19						3.32	13	42	−76	1

Differential effects of ‘Theory of mind’ and ‘Empathy’ processing: interaction between (‘ToM’ – ‘Physical 1’) and (‘Empathy’ – ‘Physical 2’)

In this comparison, neural responses in the (‘Empathy’ – ‘Physical 2’) comparison were contrasted with those in the (‘ToM’ – ‘Physical 1’) comparison to reveal areas of increased signal associated with ToM or empathy processing.

Areas more activated in the processing of ToM than empathy included right superior temporal lobe, left middle frontal gyrus, right lateral orbitofrontal cortex and cuneus (Table 2, Fig. 3). Areas which were significantly more activated in the (‘Emp’ – ‘Physical 2’) compared to the (‘ToM’ – ‘Physical 1’) contrast included right

medial prefrontal cortex, left amygdala/parahippocampal gyrus and anterior cingulate, right posterior cingulate and middle occipital gyrus (Table 2, Fig. 4).

Discussion

In this study, we investigated the neuronal networks associated with ToM and empathy. Our study differs from previously published experiments in using visual ToM and empathy stimuli in one task, thereby allowing the direct comparison of their neuroanatomy.

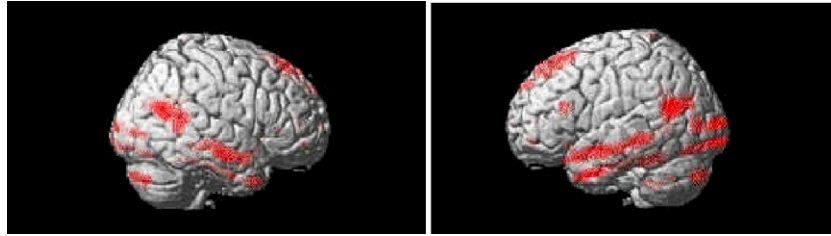


Fig. 2. Areas of activation in the conjunction analysis ('ToM' – 'Physical 1') and ('Emp' – 'Physical 2'). In this surface view, significant regions of activation are projected onto the right and left cortical surface of an average brain.

Common activations of ToM and empathy

Activations in prefrontal as well as temporal brain areas, namely medial prefrontal cortex (mPFC; Brodmann 8), temporo-parietal junction (TPJ; Brodmann 22/39) and middle and inferior temporal gyri including left temporal pole (TP; Brodmann 38) were common to both ToM and empathy. Some of these areas were identified as common activations despite the absence of a significant response in individual contrasts. This can be explained by a change in the degrees of freedom in the conjunction analysis compared to individual contrasts (i.e. 24 compared to 12) as a result of which areas which are subthreshold in individual contrasts can reach threshold in the conjunction analysis.

The medial prefrontal cortex (mPFC) has been described in imaging studies using ToM or empathy stimuli irrespective of stimulus modality (see Gallagher and Frith, 2003 for a review). Typically, activations are reported in mPFC adjacent to anterior cingulate (paracingulate) while we found significant activations in a more dorsal part of the mPFC in BA 8. This area corresponds with results by Blakemore et al. (2003) using a task requiring specific attention to the contingent nature of the relationship between animate stimuli. Blakemore et al. (2003) proposed contingency and animacy as two cues to agency underlying ToM processing. One explanation for the lack of activation in the paracingulate cortex in the current study can be derived from Walter et al. (2004). They used visual stimuli depicting cartoon stories with varying degrees of social interaction in their fMRI study; study participants had to choose the correct story ending. The 'private intention' (ToM) condition corresponds to our ToM condition in that only one character was displayed who was involved in a nonsocial action; the 'prospective social interaction' condition showed a character about to get involved in social interaction and the 'communicative intention' cartoons displayed two characters interacting with each

other. Paracingulate activation was only identified in the two latter conditions while it was not observed in the ToM condition. The authors concluded that this region was particularly sensitive to interacting mental states, but was not be associated with mentalizing per se. In the current study, a small cluster of activation in the paracingulate cortex was found in the (Emp – Ph2) – (ToM – Ph1) contrast but not in the conjunction analysis. This is in agreement with the Walter et al. (2004) interpretation of a particular role of this structure in social interaction.

The superior temporal sulcus (STS) has also consistently been described in ToM studies whereas in empathy imaging studies it has only been detected when using visual stimuli (Carr et al., 2003; Moll et al., 2002). The STS has further been associated with the perception of biological motion (Allison et al., 2000), gaze direction (Wicker et al., 1998) and attention to facial emotion (Narumoto et al., 2001). This region therefore seems to be involved in the initial analysis of social cues and the detection of intentional activity. The nearby temporo-parietal junction (TPJ) was reported by Saxe and Kanwisher (2003) in a 'false belief' task. The authors proposed a role of this region in understanding other people specific to reasoning about mental states. Temporal pole activations have been described in sentence coherence (Maguire et al., 1999), semantic decision making (Vandenberghe et al., 1996) and memory retrieval (Fink et al., 1996). Within the context of mentalizing, Frith and Frith (2003) concluded that the temporal poles are concerned with generating a wider semantic and emotional context from past experience. In empathy processing, the temporal poles have only been described in one study (Carr et al., 2003) whereas other areas of the emotional network such as the amygdala and insula seem to have more prominent roles (Carr et al., 2003; Moll et al., 2002).

Further areas of common activation between 'ToM' and 'Empathy' included orbitofrontal cortex, fusiform gyrus and

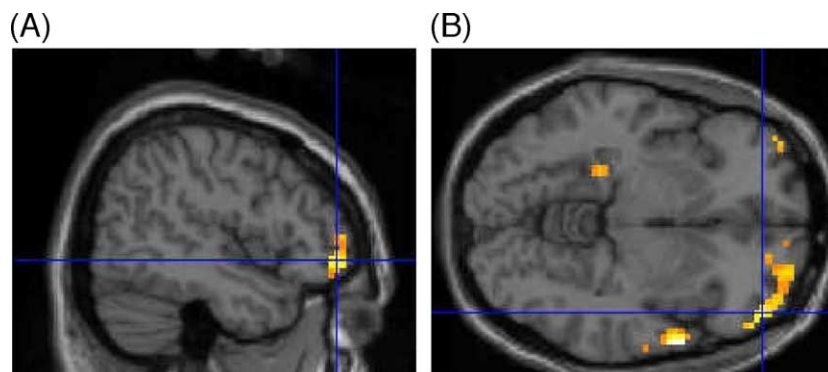


Fig. 3. Area of activation in the orbitofrontal cortex in the contrast ('ToM' – 'Physical 1') – ('Emp' – 'Physical 2'). This area was activated to a greater extent when subjects were processing ToM stimuli than during attending to empathy stimuli. (A) Sagittal view. (B) Transverse view. Crosshairs at (48; 48; -6).

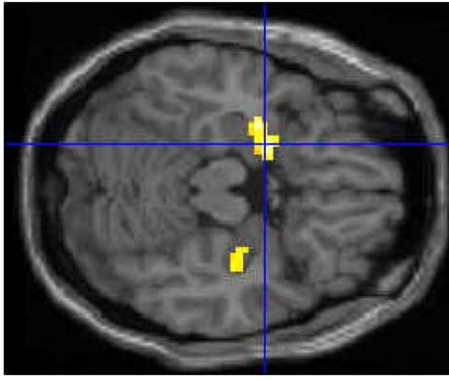


Fig. 4. Area of activation in the left amygdala in the contrast ('Emp' – 'Physical 2') – ('ToM' – 'Physical 1'). This area was activated to a greater extent when subjects were processing empathy stimuli than during ToM processing. Crosshairs at (–21; –1; –15).

lingual gyrus. The orbitofrontal cortex was highlighted in two previous ToM imaging studies (Baron-Cohen et al., 1994; Brunet et al., 2000). In empathy studies, activation of the right medial OFC was found by Farrow et al. (2001) and Moll et al. (2002). OFC lesions have been associated with disturbances in social and emotional judgment and behavior (Berlin et al., 2004). The fusiform gyrus has been related to face and object perception (Bly and Kosslyn, 1997) and has been described in a number of ToM (Castelli et al., 2000; Gallagher et al., 2000) as well as empathy imaging studies (Moll et al., 2002) particularly when visual stimuli were used. Increased lingual gyrus activation can be interpreted in terms of enhanced salience in ToM and empathy conditions (Narumoto et al., 2000).

Differential activation of ToM and empathy

Compared to ('Empathy' – 'Physical 2'), the contrast ('ToM' – 'Physical 1') revealed increased activations in lateral orbitofrontal cortex, middle frontal gyrus and superior temporal gyrus. Most of these activations were right-sided. A particular role for the right hemisphere in mentalizing is supported by neurological studies showing impairments on ToM following right hemisphere strokes (Happé et al., 1999). Furthermore, McCagh et al. (2004) reported deficits in ToM in people with focal epilepsy of right prefrontal cortex origin. The superior temporal sulcus has been suggested to have a particular role in understanding causality and intentionality (Frith, 2001a,b) of agents and it is conceivable that this function was required to a larger extent in the 'ToM' than in the 'Empathy' condition. The middle frontal gyrus activation found in the contrast between 'ToM' and 'Empathy' was located in the frontopolar cortex (BA 10), an area previously reported in tasks of executive function (Wager and Smith, 2003). We are uncertain as to an interpretation of the observed increased orbitofrontal activations in the ToM condition. This region has been associated with norm violations (Berthoz et al., 2002). A number of our ToM cartoons involved the unexpected use of objects such as the use of a broom as an oar, and one could speculate that these norm violations might account for enhanced lateral OFC activations.

Compared to ('ToM' – 'Physical 1'), the contrast ('Empathy' – 'Physical 2') revealed increased activations in left amygdala, anterior and posterior cingulate and medial prefrontal cortex. Differential activation of amygdala in the 'Empathy' condition is compatible with Blair's proposal of mediation of emotions evoked

by empathy in the amygdala. He suggests that impaired development of the amygdala may be a fundamental mechanism of failure of empathy, a core feature of psychopathy (Blair, 2003). The posterior cingulate has been found to play a role in emotional (Maddock et al., 2003) and visuospatial processing (Van Horn et al., 1998). Activations in cingulate cortex could therefore be explained by the increased emotional content of our 'Empathy' condition. The mPFC has been implicated in social perception in a range of different studies. Mitchell et al. (2002) in an fMRI study found mPFC activations to be associated with person judgments while deactivations were found when object judgments were made. The mPFC has also been associated with tasks involving self-monitoring such as perception of one's own thoughts (McGuire et al., 1996), emotions (Lane et al., 1997) and physical sensations (Blakemore et al., 1998). Kelley et al. (2002), in an event-related fMRI study, have shown selective engagement during self-referential processing, while Macrae et al. (2004) proposed that the recruitment of this region may explain memory enhancement of materials that trigger self-referential mental activity. The mPFC, particularly the paracingulate region, has also typically been associated with ToM in previous imaging studies including Brunet et al.'s (2000) from which we derived our ToM stimuli. It is therefore interesting that we failed to identify activations in this area during our ToM condition but detected an association with the processing of empathic stimuli. One difference between our and the Brunet et al.'s (2000) PET study is the use of different imaging techniques but it is difficult to see how this could fully account for the differences in activations found. However, as outlined above, other authors using an almost identical design in their fMRI study have also failed to detect paracingulate activation in mentalizing where no interaction or prospected interaction of characters was present in the stimulus material (Walter et al., 2004). It would appear that the role of the paracingulate in different types of mentalizing has to be further addressed and clarified in future research. Walter et al. (2004) suggested a particular role of paracingulate in mentalizing involving social interaction. Our results further suggest an important role of this structure in the processing of empathy provoking stimuli.

Certain limitations of our study have to be considered. In order to justify our conclusion of distinct but overlapping brain activations associated with ToM and empathy, we have to feel confident that our activation paradigm reliably isolates cognitive mentalizing and empathetic understanding in the corresponding task blocks. We used ToM stimuli which have previously been shown to be associated with a pattern of brain activation distinct from that observed in control conditions, a result which was confirmed by a principle components analysis by Brunet et al. (2000). Our 'Empathy' stimuli were piloted and shown to be easy to understand and to provoke empathic understanding of the main story character. Control conditions were matched for number of people displayed in order to avoid confounding factors of social interaction and complexity. We therefore propose that our task conditions evoked the intended mental processes in our study volunteers, and that the contrasts performed in the image analysis isolate the two variables ToM and empathy. One further possible limitation could be that one of the control conditions, the 'Physical causality two characters' condition, seemed to have been more difficult to interpret compared to 'Empathy' as judged by increased error rates. It could therefore be argued that enhanced activations during the processing of 'Empathy' compared to this control condition were due to increased coherence of the story lines in the

former condition. However, this would not account for the engagement of emotional networks in the ‘Empathy’ condition. It could further be suggested that some of our ‘Empathy’ stories depicted social rule following rather than empathy. We suggest that empathetic responding partly underlies socially desirable behavior and therefore the two processes are difficult to separate. By using a control condition involving human characters rather than physical objects, we tried to match for social perception. Furthermore, our instructions emphasized the feelings of the main character, thus facilitating the participant’s empathetic responding. The activations found involving emotional networks suggest that our task elicited such responses rather than purely social perception.

In summary, from our evidence, we suggest that both, ToM and empathy, depend on the activation of similar brain networks involved in social perception, namely the mPFC, superior temporal lobe and temporal pole. These areas form the basis for making inferences about the mental states of others. However, the appreciation of the other’s emotional states requires the additional engagement of emotional networks, particularly the amygdala. Further research is needed to advance our understanding of the interaction between these networks as well as their impairment in certain patient groups.

Acknowledgment

This research was supported by the Medical Research Council.

References

- Adolphs, R., 2003. Cognitive neuroscience of human social behaviour. *Nat. Rev., Neurosci.* 4, 165–178.
- Allison, T., Puce, A., McCarthy, G., 2000. Social perception from visual cues; role of the STS region. *Trends Cogn. Sci.* 4, 267–278.
- Ammons, R.B., Ammons, C.H., 1962. Quick Test. Missoula: Psychological Test Specialists.
- Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., Ell, P., 1994. Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults. *Br. J. Psychiatry* 165, 640–649.
- Baron-Cohen, S., Ring, H.A., Wheelwright, S., Bullmore, E.T., Brammer, M.J., Simmons, A., Williams, S.C., 1999. Social intelligence in the normal and autistic brain: an fMRI study. *Eur. J. Neurosci.* 11, 1891–1898.
- Berlin, H.A., Rolls, E.T., Kischka, U., 2004. Impulsivity, time perception, emotion and reinforcement sensitivity in patients with orbitofrontal cortex lesions. *Brain* 127, 1108–1126.
- Berthoz, S., Armony, J.L., Blair, R.J.R., Dolan, R.J., 2002. An fMRI study of intentional and unintentional (embarrassing) violations of social norms. *Brain* 125, 1696–1708.
- Blair, R.J., 2003. Neurobiological basis of psychopathy. *Br. J. Psychiatry* 182, 5–7.
- Blair, J., Sellars, C., Strickland, I., Clark, F., Williams, A., Smith, M., Jones, L., 1996. Theory of mind in the psychopath. *J. Forensic Psychiatry* 7, 15–25.
- Blakemore, S.J., Wolpert, D., Frith, C.D., 1998. Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640.
- Blakemore, S.J., Boyer, P., Pachot-Clouard, M., Meltzoff, A., Segebarth, C., Decety, J., 2003. The detection of contingency and animacy from simple animations in the human brain. *Cereb. Cortex* 13, 837–844.
- Bly, B.M., Kosslyn, S.M., 1997. Functional anatomy of object recognition in humans: evidence from positron emission tomography and functional magnetic resonance imaging. *Curr. Opin. Neurol.* 10, 5–9.
- Brunet, E., Sarfati, Y., Hardy-Bayle, M.C., Decety, J., 2000. A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage* 11, 157–166.
- Carr, L., Iacoboni, M., Dubeau, M.C., Mazziotta, J.C., Lenzi, G.L., 2003. Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. *Proc. Natl. Acad. Sci. U. S. A.* 100, 5497–5502.
- Castelli, F., Happe, F., Frith, U., Frith, C., 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* 12, 314–325.
- Davies, M., Stone, T. (Eds.), *Folk Psychology*. Blackwell, Oxford.
- Farrow, T.F., Zheng, Y., Wilkinson, I.D., Spence, S.A., Deakin, J.F., Tarrier, N., Griffiths, P.D., Woodruff, P.W., 2001. Investigating the functional anatomy of empathy and forgiveness. *NeuroReport* 12, 2433–2438.
- Ferstl, E.C., von Cramon, D.Y., 2002. What does the frontomedian cortex contribute to language processing: coherence or theory of mind? *NeuroImage* 17, 1599–1612.
- Fink, G.R., Markowitsch, H.J., Reinkemeier, M., Bruckbauer, T., Kessler, J., Heiss, W.D., 1996. Cerebral representations of one’s own past: neuronal networks involved in autobiographical memory. *J. Neurosci.* 16, 4275–4282.
- Fletcher, P.C., Happe, F., Frith, U., Baker, S.C., 1995. Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition* 57, 109–121.
- Frith, C.D., 1999. Interacting minds—A biological basis. *Science* 286, 1692–1695.
- Frith, U., 2001. Mind blindness and the brain in autism. *Neuron* 32, 969–979.
- Frith, U., 2001. The biological basis of social interaction. *Curr. Dir. Psychol. Sci.* 10, 151–155.
- Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. London, Ser. B, Biol. Sci.* 358, 459–473.
- Gallagher, H.L., Frith, C.D., 2003. Functional imaging of ‘theory of mind’. *Trends Cogn. Sci.* 7, 77–83.
- Gallagher, H.L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., Frith, C.D., 2000. Reading the mind in cartoons and stories: an fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. *Neuropsychologia* 38, 11–21.
- Gallagher, H.L., Jack, A.I., Roepstorff, A., Frith, C.D., 2002. Imaging the intentional stance in a competitive game. *NeuroImage* 16, 814–821.
- Gallese, V., 2003. The roots of empathy: the shared manifold hypothesis and the neural basis of intersubjectivity. *Psychopathology* 36, 171–180.
- Goel, V., Grafman, J., Sadato, N., Hallett, M., 1995. Modeling other minds. *NeuroReport* 6, 1741–1746.
- Goldman, A., 1995. Empathy, mind and morals. In: Davies, M., Stone, T. (Eds.), *Mental Stimulation*. Blackwell, Oxford, pp. 185–208.
- Gopnik, A., Astington, J.W., 1988. Children’s understanding of representational change and its relation to the understanding of false belief and the appearance reality distinction. *Child Dev.* 59, 26–37.
- Gordon, R.M., 1986. Folk psychology as simulation. *Mind Lang.* 1, 158–171.
- Happe, F., Brownell, H., Winner, E., 1999. Acquired ‘theory of mind’ impairments following stroke. *Cognition* 70, 211–240.
- Heal, J., 1995. How to think about thinking. In: Davies, M., Stone, T. (Eds.), *Mental Stimulation*. Blackwell, Oxford, pp. 33–52.
- Kelley, W.M., Macrae, C.N., Wyland, C.L., Caglar, S., Inati, S., Heatherton, T.F., 2002. Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* 14, 794–885.
- Lane, R.D., Fink, G.R., Chua, P.M., Dolan, R.J., 1997. Neural activation during selective attention to subjective emotional responses. *NeuroReport* 8, 3969–3972.
- McCabe, K., Houser, D., Ryan, L., Smith, V., Trouard, T., 2001. A functional imaging study of cooperation in two-person reciprocal exchange. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11832–11835.
- McCagh, J., Corcoran, R., Baker, G.A., Fisk, J.E., 2004. Social cognition in epilepsy. *Proc. Br. Psychol. Soc.* 12, 70.
- McGuire, P.K., Silbersweig, D.A., Frith, C.D., 1996. Functional neuroanatomy of verbal self-monitoring. *Brain* 119, 907–917.

- Macrae, C.N., Moran, J.M., Heatherton, T.F., Banfield, J.F., Kelley, W.M., 2004. Medial prefrontal activity predicts memory for self. *Cereb. Cortex* 14, 647–654.
- Maddock, R.J., Garrett, A.S., Buonocore, M.H., 2003. Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Hum. Brain Mapp.* 18, 30–41.
- Maguire, E.A., Frith, C.D., Morris, R.G.M., 1999. The functional neuroanatomy of comprehension and memory: the importance of prior knowledge. *Brain* 122, 1839–1850.
- Mitchell, J.P., Heatherton, T.F., Macrae, C.N., 2002. Distinct neural systems subservise person and object knowledge. *Proc. Natl. Acad. Sci. U. S. A.* 99, 15238–15243.
- Moll, J., Oliveira-Souza, R., Eslinger, P.J., Bramati, I.E., Mourao-Miranda, J., Andreiuolo, P.A., Pessoa, L., 2002. The neural correlates of moral sensitivity: a functional magnetic resonance imaging investigation of basic and moral emotions. *J. Neurosci.* 22, 2730–2736.
- Narumoto, J., Yamada, H., Iidaka, T., Sadato, N., Fukui, K., Itoh, H., Yonekura, Y., 2000. Brain regions involved in verbal or non-verbal aspects of facial emotion recognition. *NeuroReport* 11, 2571–2576.
- Narumoto, J., Okada, T., Sadato, N., Fukui, K., Yonekura, Y., 2001. Attention to emotion modulates fMRI activity in human right superior temporal sulcus. *Cogn. Brain Res.* 12, 225–231.
- Premack, D., Woodruff, G., 1978. Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526.
- Ruby, P., Decety, J., 2004. How would you feel versus how do you think she would feel? A neuroimaging study on perspective-taking with social emotions. *J. Cogn. Neurosci.* 16, 988–999.
- Russell, T.A., Rubia, K., Bullmore, E.T., Soni, W., Suckling, J., Brammer, M.J., Simmons, A., Williams, S.C., Sharma, T., 2000. Exploring the social brain in schizophrenia: left prefrontal underactivation during mental state attribution. *Am. J. Psychiatry* 157, 2040–2042.
- Sarfati, Y., Hardy-Bayle, M.C., Besche, C., Widlocher, D., 1997. Attribution of intentions to others in people with schizophrenia: a non-verbal exploration with comic strips. *Schizophr. Res.* 25, 199–209.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking people: the role of the temporo-parietal junction in “theory of mind”. *NeuroImage* 19, 1835–1842.
- Sheehan, D.V., Lecrubier, Y., Harnett-Sheehan, K., Janavs, J., Weiller, E., Bonara, L.I., Keskiner, A., Schinka, J., Knapp, E., Sheehan, M.F., Dunbar, G.C., 1997. The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *Eur. Psychiatr.* 12, 224–231.
- Singer, T., Seymour, B., O’Doherty, J., Kaube, H., Dolan, R.J., Frith, C.D., 2004. Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162.
- Skuse, D.H., James, R.S., Bishop, D.V.M., Coppin, B., Dalton, P., Aamodt-Leeper, G., Bacarese-Hamilton, M., Creswell, C., McGurk, R., Jacobs, A., 1997. Evidence from Turner’s syndrome of an imprinted X-linked locus affecting cognitive function. *Nature* 387, 705–708.
- Soderstrom, H., 2003. Psychopathy as a disorder of empathy. *Eur. Child Adolesc. Psychiatry* 12, 249–252.
- Stone, V.E., Baron-Cohen, S., Calder, A., Keane, J., Young, A., 2003. Acquired theory of mind impairments in individuals with bilateral amygdala lesions. *Neuropsychologia* 41, 209–220.
- Stuss, D.T., Gallup, G.G., Alexander, M.P., 2001. The frontal lobes are necessary for ‘theory of mind’. *Brain* 124, 279–286.
- Talairach, J., Tournoux, J., 1998. *Coplanar stereotactic atlas of the human brain*. Thieme Verlag, Stuttgart.
- Vandenberghe, R., Price, C., Wise, R., Josephs, O., Frackowiak, R.S.J., 1996. Functional anatomy of a common semantic system for words and pictures. *Nature* 383, 254–256.
- Van Horn, J.D., Gold, J.M., Esposito, G., Ostrem, J.L., Mattay, V., Weinberger, D.R., Berman, K.F., 1998. Changing patterns of brain activation during maze learning. *Brain Res.* 793, 29–38.
- Vogeley, K., Bussfeld, P., Newen, A., Herrmann, S., Happe, F., Falkai, P., Maier, W., Shah, N.J., Fink, G.R., Zilles, K., 2001. Mind reading: neural mechanisms of theory of mind and self-perspective. *NeuroImage* 14, 170–181.
- Walter, H., Adenzato, M., Ciaromidaro, A., 2004. Understanding intentions in social interaction: the role of the anterior paracingulate cortex. *J. Cogn. Neurosci.* 16, 1854–1863.
- Wager, T.D., Smith, E.E., 2003. Neuroimaging studies of working memory: a meta-analysis. *Cogn. Affect. Behav. Neurosci.* 3, 255–274.
- Wicker, B., Michel, F., Henaff, M.-A., Decety, J., 1998. Brain regions involved in the perception of gaze: a pet study. *NeuroImage* 8, 221–227.
- Williams, J.H., Whiten, A., Suddendorf, T., Perrett, D.I., 2001. Imitation, mirror neurons and autism. *Neurosci. Biobehav. Rev.* 25, 287–295.