doi:10.1093/scan/nsn047 SCAN (2009) 4, 166–176

Dysfunction of the social brain in schizophrenia is modulated by intention type: An fMRI study

Henrik Walter,[1] Angela Ciaramidaro,[2] Mauro Adenzato,[2] Nenad Vasic,[3] Rita Bianca Ardito,[2] Susanne Erk,[1] and Bruno G. Bara[2]

[1]Department of Psychiatry and Psychotherapy, Division of Medical Psychology, University of Bonn, Germany, [2]Department of Psychology, Center for Cognitive Science, University of Turin, Italy, and [3]Department of Psychiatry, University of Ulm, Germany

In this fMRI study, we investigated theory of mind (ToM) in patients with paranoid schizophrenia. We hypothesized that the network supporting the representation of intentions is dysfunctional in patients with schizophrenia dependent on the type of intention involved. We used a paradigm including a control condition (physical causation) and three intention conditions (private intention, prospective social intention and communicative intentions) differing in the degree of social interaction. In all four experimental conditions patients performed worse than controls regarding accuracy and reaction time. They showed significantly less activation in three regions typically activated in ToM tasks, i.e. paracingulate cortex and bilateral temporo-parietal junctions. However, this dysfunction was dependent on the type of intention represented, i.e. was present only for social but not for non-social intentions. Moreover, part of the reduced activation was related to the fact that there was no signal drop in these regions for the physical causality condition as usually found in controls. This may be due to the tendency of schizophrenic patients to attribute intentionality to physical objects. Our findings have implications for the study and understanding of ToM in schizophrenia but also in other disorders like autism.

Keywords: schizophrenia; theory of mind; social interaction; communication; intention

INTRODUCTION

Theory of mind (ToM), sometimes also referred to as 'mind reading' or 'mentalizing', is the cognitive ability to understand others as intentional agents by inferring and thereby representing their mental states (Premack and Woodruff, 1978; Baron-Cohen, 1995; Frith, 2004). From an evolutionary point of view such an ability is a selective adaptation to an increasingly complex social environment helping to explain and predict behaviour in others and is linked to successful social interaction (Brothers, 1990; Dunbar, 1998). Functional imaging studies in healthy controls have shown four key regions involved in ToM: The medial prefrontal cortex (MPFC), the posterior cingulate cortex/precuneus and bilateral temporo-parietal regions (Gallagher and Frith, 2003; Saxe, 2006; Ciaramidaro *et al.*, 2007), which we will call the ToM network. The prominent view is that the most specific structure for mentalizing is a subregion of the MPFC, the paracingulate cortex (Gallagher and Frith, 2003; Amodio and Frith, 2006; Frith and Frith, 2006), although it has recently been argued that this role should be attributed to the right temporo-parietal junction (TPJ) (Saxe, 2006). These apparently inconsistent findings can be explained by the fact that different nodes of the ToM network are modulated by different types of intentional states (Ciaramidaro *et al.*, 2007): Using cartoons we found activation of the posterior cingulate cortex and right TPJ suffices to solve tasks involving private intentions (PInt), i.e. intentions of a single agent directed at objects, whereas the MPFC is necessary when social interaction is *prospective* (PSInt), i.e. intention of one agent (A) preparing to interact with B (who is not actually present), or *here and now*, i.e. communicative intention (CInt) (Bara, 2008) involving two people actually interacting. This last condition is also the only one that recruited all four areas described above (Walter *et al.*, 2004).

Impairment in ToM has been implicated in various psychiatric disorders like autism (Baron-Cohen, 1995), frontotemporal dementia (Gregory *et al.*, 2002), depression (Hynes *et al.*, 2006), brain lesions of the ToM-network (Rowe and Passingham, 2001; Stuss *et al.*, 2001; Stone *et al.*, 2003; Apperly *et al.*, 2004) and schizophrenia (Brüne, 2005). Schizophrenia is a heterogeneous disorder with various symptom subgroups showing different degrees of mentalizing impairment and this deficit seems to depend on their psychopathology (disorganized > paranoid > patients with passivity phenomena) (Corcoran *et al.*, 1997; Pickup and Frith, 2001; Sprong *et al.*, 2007). Frith (1992) proposed that certain psychotic symptoms associated with schizophrenia reflect a deficit in the ability of mentalizing and claims that this is the result of a failure of patients to monitor their own and others' mental states and behaviour.

Received 23 July 2008; Accepted 25 November 2008
Advance Access publication 14 March 2009

This work was supported by Regione Piemonte (Bando regionale per la ricerca scientifica applicata, cod. A239).

Correspondence should be addressed to Henrik Walter, MD, PhD, Department of Psychiatry, Division of Medical Psychology, University of Bonn, Sigmund-Freud-Straße 25, D-53105 Bonn, Germany. E-mail: henrik.walter@ukb.uni-bonn.de.

This deficit might have different reasons: While disorganized patients lack the basic capability to represent mental states at all, patients with paranoid symptoms still have this competence but use it inappropriately.

However, not only reduced ToM ability is maladaptive. It has been argued that an exaggerated ToM also has costs and can be maladaptive (Brüne, 2001), as may be the case in schizophrenia (Abu-Akel and Bailey, 2000; Frith, 2004). Indeed, in humans, the ToM-module is apparently permanently 'online' screening other persons, sometimes even non-living objects, for their putative intentions. Therefore, humans are always at risk of making false conjectures about other people's intentions or to anthropomorphize actions of other beings or non-living events by assuming intentions where in fact there are none (Brüne, 2005). Whereas healthy persons are able to reflect on the appropriateness and correctness of these more or less automatic attributions, patients with paranoid schizophrenia might over-attribute significance and intentions to events, person and objects. Thus, paranoid patients may be characterized by hyper-intentionality. Abu-Akel and Bailey (2000) speak about 'hyper ToM', an attitude associated with quantitative over-generation of hypotheses or over-attribution of mental states. Paranoid patients seem to have an intact ToM in the sense that they experience others as possessing mental states, but they are impaired in using contextual information inducing them to make incorrect 'online' inferences about others (Brüne, 2005). One promising way to understand the nature of ToM deficits in schizophrenia is to investigate the activation patterns of the ToM network with functional neuroimaging.

To date, only few studies have investigated ToM tasks in patients with schizophrenia using emotion labelling (Russell *et al.*, 2000), cartoons (Brunet *et al.*, 2003; Brüne *et al.*, 2008), and empathy and forgiveness judgements (Lee *et al.*, 2006). These studies yielded inconsistent results with hypo- (Russel *et al.*, 2000; Brunet *et al.*, 2003; Lee *et al.*, 2006) as well as hyper-activation (Brüne *et al.*, 2008) of nodes of the ToM network, in particular in the medial prefrontal cortex. Moreover, three of these studies used less than 10 (5, 7 or 9) subjects and none of them differentiated between different types of intentional states.

Based on the idea that ToM is a function of the social mind and that patients with schizophrenia are impaired in social functioning, we were interested to investigate brain activation in a homogeneous group with paranoid schizophrenia for non-social and social ToM tasks with a validated task (Walter *et al.*, 2004; Ciaramidaro *et al.*, 2007) using three different types of intention (private intention, prospective social intention and communicative intention) and a physical causation control condition. We hypothesized to find a dysfunction in the mentalizing network in terms of reduced brain activations in the intentional conditions, in particular for communicative intentions (CInt), because the schizophrenic patients' attitude of 'over-attributing' intentions seems to be related to violations of pragmatic rules in their use of language and incorrect inferences of communicative intentions (Brüne, 2005).

METHODS

Subjects

We studied 14 right-handed patients (seven females) with paranoid schizophrenia according to ICD-10 (F 20.0)/DSM-IV, recruited from among the inpatients treated at the Department of Psychiatry at the University of Ulm, as well as a matched control group (Table 1). Patients diagnosed with concurrent axis I disorder according to DSM-IV criteria were excluded from the study. In addition to a detailed interview conducted by an experienced clinical psychiatrist (N.V.), case notes were reviewed to corroborate the diagnosis of DSM-IV schizophrenia. Psychopathology was rated by means of the Brief Psychiatric Rating Scale (BPRS) (Overall and Gorham, 1976). In order to provide additional information on negative symptoms (Andreasen *et al.*, 2005), the BPRS was complemented by the negative symptoms subscale, as measured by the Positive and Negative Syndrome Scale (PANSS) (Kay *et al.*, 1989). To minimize cerebral activation effects associated with poor performance, we included only subjects with an accuracy clearly above chance during performance of the fMRI activation task (at least 60% correct responses, the chance level was 33%). This criterion led us to the exclusion of two patients (one female), and we included 12 subjects (six females) with mean age 29.41 years (s.d. 5.96) in the functional analysis. The healthy control group consisted of 12 right-handed subjects (six females), mean age 24.75 years (s.d. 2.63). Exclusion criteria for the control group were presence or history of a neurological or psychiatric disorder, substance abuse or dependence. The University of Ulm's ethics committee approved the study. Prior to scanning, written informed consent was obtained from all subjects according to the Declaration of Helsinki.

Experimental design

In our task, participants were asked to look at short comic strips and then choose a picture that showed the only logical

Table 1 Demographic and clinical characteristics of the sample

	Patients ($n = 12$) Mean (s.d.)	Control Subjects ($n = 12$) Mean (s.d.)	SZ versus C *P*
Age (year)	29.5 (6.0]	24.75 (2.6)	0.02
Male/Female	6/6	6/6	
Education (years)	12.4 (1.1]	12.3 (1.4)	0.73
Positive and Negative Syndrome Scale			
Positive	17.75 (5.06)	–	
Negative	19.41 (3.96)	–	
Total (positive, negative, general)	73.75 (11.02)	–	
Brief Psychiatric Rating Scale	47.18 (10.6)	–	
HAMD	12.18 (4.57)	–	
Years of Illness	6.3 (5.2)	–	

story ending. Comic strips (examples available at www.psych.unito.it/csc/pers/adenzato/pdf/intention_protocol.pdf) pertained to the following experimental categories: (i) PInt: Private intention. In this condition, participants represented another person's intention, based on observation of that person's isolated action, e.g. observing a single person (A) changing a broken bulb in order to read a book; (ii) PSInt: Prospective social intention (potentially shared in the future). In this condition, participants represented another person's intention to socially interact with someone else in the future, based on the observation of that person's isolated action, e.g. observing a single person (A) preparing a romantic dinner for another person (B), who is not yet present in the scenario; (iii) CInt: Communicative intention (shared in the present). In this condition, participants represented the intentions to communicate based on the observation of two people interacting, e.g. observing a person (A) obtaining a glass of water by asking another person (B) to get it for her. The control condition was physical causality (Ph-C), in which participants represented non-intentional causal links among objects, e.g. a ball blown by a gust of wind knocking over and breaking a glass of water.

We presented comic strips consisting of a sequence of three pictures (the story-phase); each picture was displayed for 3 s (Figure 1). The story phase was followed by a choice-phase, during which three possible solutions were displayed simultaneously for 7 s. Thus, one trial (one comic strip) lasted 16 s (story-phase plus choice-phase). The participants' task was to choose the logical story ending. Participants indicated their choice by pressing one of three buttons as quickly as possible. Only one picture represented the correct answer. The two uncorrected pictures were constructed according to the following principle: One foil picture showed a possible, but illogical ending. The second foil included the objects of the last scene rearranged physically without containing a real action. We used a slow event-related design with a relatively long inter-trial interval (rest period) of 7–11 s (jittered) between trials. Eleven comic strips were presented for each of the four conditions, summing up to a total of 44 trials. The comic strips and the visual location of the correct answer were presented in pseudo randomized order. The experimental protocol was administered in two sessions of 22 trials each. Before scanning each participant received training with additional comic strips for each category in order to verify that the subject had clearly understood the instructions. During scanning, participants wore luminescent crystal display glasses ('goggles'; Resonance Technologies, Northridge, CA). Stimuli were presented with Presentation software (Neurobehavioral Systems).

Behavioral data analysis. Participant reaction times and response accuracy were measured during scanning. Data were analyzed in a one-way ANOVA with subsequent comparisons between means, using Bonferroni's *post hoc* test. Psychopathology of patients and controls was measured using the BPRS and PANNS scale. To determine possible relations between medication (chlorpromazine equivalent), negative and positive schizophrenia symptoms according to PANNS, and behavioural performance in the ToM task (performance, reaction times) on the one hand and activation patterns on the other, we correlated the first eigenvariate from the peak voxel of each significantly activated cluster with the respective scores. Statistical analyses were carried out using SPSS 11.5 for Windows.

fMRI data acquisition and analysis. fMRI data were acquired using a 1.5 Tesla Siemens Magnetom Symphony whole-body MRI System equipped with a head volume coil. T_2*-weighted functional MR images were obtained using echo-planar imaging in an axial orientation. Image size was 64×64 pixels, with a field of view of 192 mm. One volume covering the whole brain consisted of 25 slices with 4 mm slices thickness and a 1 mm gap. Time of repetition (TR) was 2.250 s, echo time (TE) was 40 ms.

One session contained 257 volumes. The first four volumes of each session were discarded in order to allow for T_2-equilibration. Data pre-processing and statistical analysis were conducted with SPM 2 (Statistical Parametric Mapping, Wellcome Institute of Cognitive Neurology, London, UK) and MATLAB 6.3 (MathWorks, Natick, Massachusetts, USA) using standardised procedures

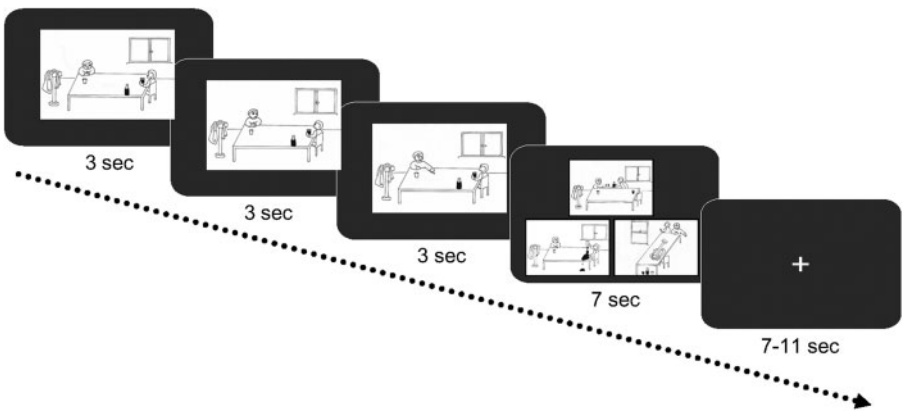

Fig. 1 Activation paradigm, exemplarily shown for a trial of the CInt condition.

(Friston *et al.*, 1995). Individual functional images were slice-timed and corrected for motion artefacts starting from the fifth volume. Images were spatially normalized ($3 \times 3 \times 3\,mm^3$) using sinc interpolation to the echo-planar template of SPM in MNI space and spatially smoothed with an 8 mm full width at half maximum isotropic Gaussian kernel. The variance of each voxel was estimated for each condition according to the general linear model. Images were globally scaled; high frequency noise was removed using a low pass filter (Gaussian kernel with 4.0 s FWHM); and low frequency drifts were removed via a high pass filter. Analyses were performed on two levels.

In a first-level analysis each subject was analysed separately. Regressors were defined for story-phase and choice-phase for each of the four conditions separately as box cars convolved with the canonical hemodynamic response function implemented in SPM2. Realignment parameters were included in the model. Contrast images for each condition (Ph-C, PInt, PSInt and CInt) were calculated by using the regressors for story and choice phases together. To account for interindividual variance and in order to generalise inferences (Holmes and Friston, 1998), we conducted a second-level analysis.

For within-group comparisons we calculated an analyses of variance (ANOVA) on the second level and contrasted each of the intentional conditions (PInt, PSInt and CInt) with the control condition (Ph-C). Group comparisons between control subjects and patients with schizophrenia were also computed using an ANOVA and we used the same contrasts utilised in the within-group comparisons.

In order to exclude that the resulting activation was not due to the different behavioural results (patients made more errors and were slower than the control group, see also the 'behavioural results' section), we calculated three different models for the group comparisons between control subjects and patients with schizophrenia: *Model 1*: For single subject (first level) analyses, all trials were included, i.e. correct and incorrect (for control and patient group, respectively). *Model 2*: For single subject (first level) analyses, only correct trials were included, i.e. incorrect and omitted trials were removed. These trials were pooled and used as individual regressors of no interest for each subject. In this case, the resulting model is composed of more trials for the control group than for the patient group (remember that patients made more errors). *Model 3*: For single-subject (first-level) analyses, we included for the control group the same number of correct trials as for the patients' group. That means that we had to exclude some trials of the control groups. These trials were also pooled and used as individual regressors of no interest for each subject. Both first and second-level analyses were performed in exactly the same way for Models 2 and 3 as in Model 1. As all regions included a priori defined regions of interests (Walter *et al.*, 2004; Ciaramidaro *et al.*, 2007) we chose an uncorrected threshold of $P < 0.001$ at the voxel level, corrected for extent ($P < 0.05$) at the cluster level (Forman *et al.*, 1995). Anatomical regions and denominations are reported according to the atlases of Talairach and Tournoux (1988) and Duvernoy (1999). Coordinates are maxima in a given cluster according to the standard MNI-template.

RESULTS

Behavioral results

Analysis of reaction times in milliseconds (for correct answers only) showed a main effect of group: Patients were slower than the control group $F(1,22) = 9.105$, $P = 0.006$ but there was no interaction between reaction time and group [$F(3,66) = 0.593$, $P = 0.62$]. Similar results were obtained for response accuracy: patients made more errors than the control group [$F(1,22) = 37.8$, $P < 0.0001$] but there was no interaction between condition and group [$F(3,66) = 0.39$, $P = 0.75$]. Non-parametric direct comparisons between patients and control for each condition using t-tests showed that patients were slower and made more errors in each condition. *T*-tests for reaction time revealed the following results: For PhC $P = 0.026$, for PInt $P = 0.013$, for PSInt $P = 0.011$ and CInt $P = 0.002$; *T*-tests for accuracy for all four comparisons $P > 0.001$.

Neuroimaging results

Main functional data analysis, including all performed trials (Model 1). Separate random effects analyses of activated regions comparing each experimental condition (PInt, PSInt and CInt) with the control condition (Ph-C) in schizophrenia patients and the control group showed significant differences in activation patterns within and between groups. The group activations of linear contrast for controls and patients showed partially similar activation patterns (see also Figure 2 and Table 2). However, there was less activation in patients with schizophrenia in the posterior cingulate cortex and bilateral TPJ, but nearly no activation in medial prefrontal cortex.

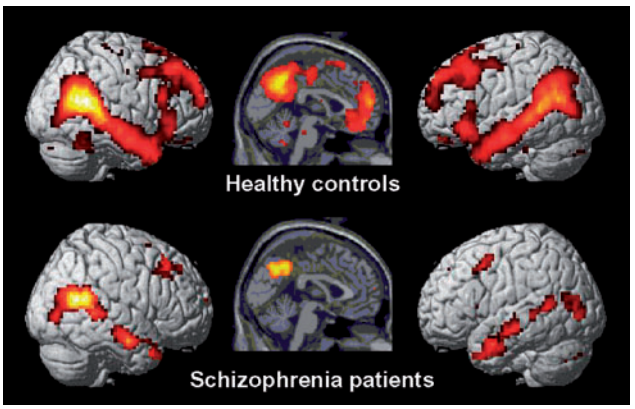

Fig. 2 Main effects for the three experimental conditions (PInt, PSInt, CInt) versus control condition (Ph-C) within-groups (Analysis of Model 1). Second level ANOVA, $P < 0.05$ FDR corrected.

Table 2 Coordinates and anatomical localization for the three experimental conditions (PInt, PSInt, CInt) versus control condition (Ph-C)

		PInt				PSInt				CInt			
Region	BA	*x*	*y*	*z*	*Z*	*x*	*y*	*z*	*Z*	*x*	*y*	*z*	*Z*
Model 1: Control group													
Precuneus	7	0	−57	45	4.28[a,b]	0	−63	30	5.57[a,b]	0	−57	45	>8[a,b]
	7	3	−63	33	3.38[a,b]	0	−57	42	5.48[b]				
Cingulate gyrus	29/30					9	−51	6	5.34[b]				
Medial occipital gyrus	19/39	48	−69	6	6.52[a,b]					48	−75	6	5.19[b]
						−48	−75	6	5.03[b]	−51	−75	6	7.67[b]
Superior temporal gyrus	22	48	−57	12	7.05[a,b]	45	−54	12	>8[a,b]	48	−54	12	>8[a,b]
										−57	−45	5	7.55[a,b]
Tempoparietal Junction	22	48	51	21	3.87[a,b]					48	−57	21	>8[a,b]
						−42	−57	21	4.88[b]	−45	−57	21	>8[a,b]
Medial temporal gyrus	21									48	−36	−3	7.55[b]
										−57	−45	6	7.55[b]
Anterior temporal pole	21					51	−9	−18	5.42[b]	54	−3	−21	7.09[b]
						−54	−6	−21	5.01[b]	−54	−3	−24	6.55[b]
Ventrolateral prefrontal cortex	47									57	27	−3	5.15[b]
	47									−48	30	−12	6.08[b]
Dorsolateral prefrontal cortex	44/46					24	36	33	4.44[b]	27	36	33	5.52[b]
										−24	33	36	5.01[b]
Superior frontal sulcus	8									12	51	39	5.04[b]
										−15	51	39	6.53[a,b]
Paracingulate cortex	10/32					6	57	6	4.39[b]	3	57	27	6.98[a,b]
						−6	57	15	4.94[a,b]				
Orbito frontal cortex	11					3	48	−12	4.88[a,b]	0	45	−21	6.51[b]
						−3	24	−21	4.21[b]				
Cerebellum						45	−45	−18	5.68	42	−51	−18	5.81[b]
Model 2: Control group													
Precuneus	7	0	−54	45	3.58	3	−60	30	4.94[b]	3	−57	30	7.80
						0	−54	42	4.34[b]				
Cingulate gyrus	29/30					9	−51	6	4.90				
Medial occipital gyrus	19/39									51	−69	6	4.72
		−48	−75	6	3.59	−48	−75	9	3.90[b]	−51	−75	6	4.71
Superior temporal gyrus	22	51	−39	12	5.53[a,b]	45	−51	9	7.28[a,b]	54	39	12	7.47
										−51	−45	6	5.04
Tempoparietal Junction	22	45	−57	15	6.11[a,b]	51	−57	21	7.10[a,b]	45	−57	21	>8[a,b]
						−42	−54	24	4.34[b]	−45	−57	21	7.53
Medial temporal gyrus	21									48	−36	−3	6.90
										−51	−36	−3	5.62
Anterior temporal pole	21					51	−6	−24	4.66[b]	57	−3	−21	7.09
						−54	−3	−24	4.24	−57	0	−21	7.08
Ventrolateral prefrontal cortex	47									57	27	−3	5.15[b]
	47									−48	30	−12	6.08[b]
Dorsolateral prefrontal cortex	44/46					27	36	39	3.97	24	39	42	5.82[b]
										−24	45	42	4.54
Superior frontal sulcus	8									12	51	39	5.04
										−15	51	33	6.49
Paracingulate cortex	10/32					6	60	18	4.20[b]	6	60	21	7.09
Orbito frontal cortex	11					3	51	−15	4.91[b]	0	51	−15	6.41
						−3	33	−21	4.03[b]				
Cerebellum		45	−42	−21	3.80	45	−45	−18	4.25				
Parahippocampal gyrus						−24	−27	−21	4.21				
Model 3: Control group													
Precuneus	7	0	−60	45	3.51	3	−60	27	4.94[b]	0	−60	33	7.31[a,b]
	7												
Cingulate gyrus	29/30					9	−51	6	3.91[b]				

continued

Table 2 Continued

		PInt				PSInt				CInt			
Region	BA	*x*	*y*	*z*	*Z*	*x*	*y*	*z*	*Z*	*x*	*y*	*z*	*Z*
Medial occipital gyrus	19/39	51	−69	6	5.15[a,b]								
		−48	−75	9	3.85[a,b]	−48	−79	9	4.50[b]	−51	−69	12	4.59[a,b]
Superior temporal gyrus	22					45	−54	9	6.50[b]	48	−51	12	5.41[a,b]
						−57	−45	12	4.12	−57	−45	3	5.89[a,b]
Tempoparietal Junction	22	48	−60	18	6.35[a]	48	−60	18	6.24[a,b]	48	−60	24	6.63[a,b]
						−39	−57	24	4.93[b]	−51	−57	21	6.55[a,b]
Medial temporal gyrus	21												
										−57	−42	3	5.89[a,b]
Anterior temporal pole	21					51	−6	−21	5.05[b]	54	−3	−21	5.34[a,b]
						−54	−3	−24	4.76[b]	−54	−3	−24	5.32[a,b]
Ventrolateral prefrontal cortex	47									54	30	−9	3.85[b]
	47									−48	−24	−15	3.90[b]
Dorsolateral prefrontal cortex	44/46					27	36	42	3.99[b]	24	35	39	4.53[b]
										−27	35	33	3.98[b]
Superior frontal sulcus	8									15	42	48	3.98[b]
										−15	51	42	5.47[b]
Paracingulate cortex	10/32					6	51	12	4.55[b]	6	57	12	5.75[b]
						−6	57	12	3.80[b]				
Orbito frontal cortex	11					3	36	−21	3.93[b]	0	33	−6	3.59[a,b]
						−6	24	−18	3.91[b]				
Cerebellum						45	−48	−18	4.93[b]				
						−42	−48	−21	3.96[b]				
Model 1: Patients group													
Medial occipital gyrus	19	45	−81	3	5.01[a,b]	45	−81	0	5.09[a,b]				
		−48	−78	0	3.84[b]	−48	−81	0	4.11[b]	−45	−81	0	3.91[b]
Superior temporal gyrus	22									45	−63	15	6.22[a,b]
										−57	−54	12	4.40[b]
Tempoparietal Junction						45	−63	18	5.38[a,b]	45	−69	21	5.87[a,b]
Medial temporal gyrus	39												
										−45	−78	18	3.51[b]
Anterior temporal pole	21					57	−9	−15	3.88	57	−9	−18	4.63[b]
Dorsolateral prefrontal cortex	8/6									−39	9	45	4.56[b]
Model 2 − 3: Patients group[a]													
Precuneus	7									0	−63	33	3.91
Medial occipital gyrus	19	48	−81	3	4.55[b]	45	−81	0	3.82				
		−48	−81	0	3.23	−45	−78	15	3.54				
Superior temporal gyrus	22									51	−36	6	4.50
						−36	−63	12	3.43	−33	−63	9	3.45
Tempoparietal Junction	22					42	−63	21	3.61	45	−69	21	4.15
Anterior temporal pole	21									57	−12	−18	4.73
										−53	−9	−18	3.41
Dorsolateral prefrontal cortex	8/6									27	21	36	3.79

Second level within-group ANOVA, $P < 0.001$ uncorrected, a = ($p < 0.05$ FWE corrected), b = ($p < 0.05$ FDR corrected).
x, y, and z are Talairach coordinates of the most significant centre of activation within a cluster; Z = z-value; BA = putative Brodmann Area.
Models 2 and 3 are identical for the patient group (patients made more errors). Only for the control group Models 2 and 3 are different.

FMRI analysis—between group effects (Model 1). ANOVA between groups comparing contrasts of interest in controls and patients revealed significantly elevated activations only in the control group compared to the patient group (Table 3) and not vice versa. They are shown in Figure 3 and Table 3. As can been seen in Figure 3 the contrast PInt versus Ph-C reveals no differences between the two groups in nodes of the ToM network. In the contrast PSInt versus Ph-C the right TPJ and MPFC were activated significantly more in the control group and the contrast CInt versus Ph-C indicates significant group differences in the right and left TPJ and the MPFC.

Table 3 Coordinates and anatomical localizations where healthy controls showed relatively more cerebral activation compared with paranoid patients for the three experimental conditions (PInt, PSInt, CInt) versus control condition (Ph-C)

		PInt*				PSInt*				CInt**			
Region	BA	*x*	*y*	*z*	*Z*	*x*	*y*	*z*	*Z*	*x*	*y*	*z*	*Z*
Model 1													
Paracingulate cortex	10/32					3	57	12	4.26	15	57	21	4.48
										−3	57	9	4.14
Tempoparietal junction	22					45	−54	9	4.26	45	−54	24	3.82
										−48	−48	24	4.55
Model 2													
Precuneus										−6	−48	24	4.43
Paracingulate cortex	10/32					21	54	18	3.73				
Tempoparietal junction	22												
										−48	−48	24	4.88
Model 3													
Superior Frontal gyrus	10					18	54	18	4.24				
						−18	48	39	4.31				
Paracingulate cortex	10/32					9	57	9	4.21				
						−6	57	9	4.02	−6	57	9	3.12
Tempoparietal junction	22					45	−54	9	4.26	48	−51	9	3.56
						−48	−48	21	3.49	−48	−48	24	4.23

Second level within-group ANOVA, $P < 0.001$ uncorrected, *for uncorrected extended cluster 19, **uncorrected, $P < 0.05$ at the cluster level.
x, y and z are Talairach coordinates of the most significant centre of activation within a cluster; Z = Z-value; BA = putative Brodmann Area.

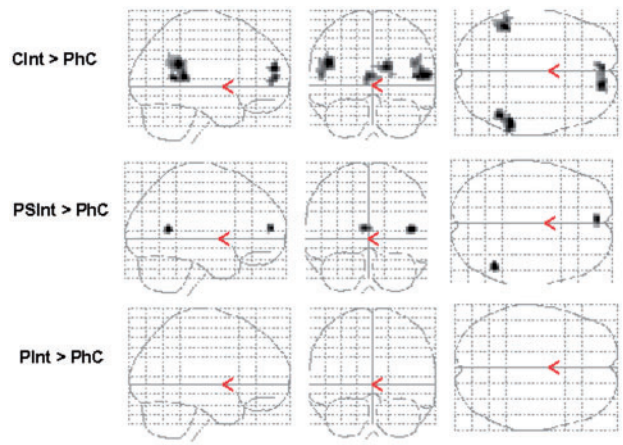

Fig. 3 Regions in which healthy controls showed relatively more cerebral activation compared with paranoid patients for the three experimental conditions (PInt, PSInt, CInt) versus control condition (Ph-C). Second level ANOVA, $P < 0.001$ uncorrected, $P < 0.05$ at the cluster level. For location within the MNI-template compare Figure 4.

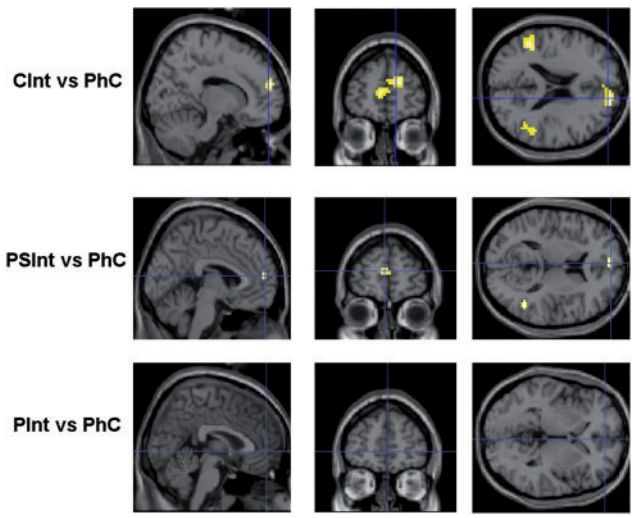

Fig. 4 Regions from Figure 3 shown in the MNI template.

A clearer picture emerges if the mean activation sizes per condition and group are plotted (compare Figure 4). Both groups exhibit an increasing linear relationship in activation in the right TPJ in the order Ph-C < PInt < PSInt < CInt. A different trend can be noticed for the left TPJ and MPFC: Whereas the control group also exhibited parametric activation in this region, the patient group showed positive beta parameters for the Ph-C and for the CInt condition and negative beta parameters that were similar for the PInt and PSInt conditions.

Functional data analysis for Models 2 and 3. In Model 2 (only correct trials) and in Model 3 (same number of correct trials for controls and patient group) we found results that are basically identical to the ones described

for Model 1 (more details in Table 3), i.e. the same within group results and the same between group differences. Therefore, we use and discuss only the results of Model 1. For between group results of all three models see Table 3.

FMRI-correlations. We could not find any significant correlations between brain activation patterns with medication (chlorpromazine equivalents), patients' positive and negative scores on the PANNS and BPRS scale or performance and reaction times.

DISCUSSION

The present work aimed at investigating the dysfunction of the ToM network in a homogenous group of patients with paranoid schizophrenia and its modulation by different intention types. We used a ToM task including three different types of intention: PInt, PSInt and CInt. Confirming our hypothesis, we found a dysfunction of the ToM network, i.e. in the medial prefrontal cortex and the right and left temporo-parietal cortex which was modulated by type of intention. For communicative intentions all three regions showed less differential activation in patients, for prospective intentions only the right TPJ and the MPFC were less activated and for private intentions there were no group differences in activation. Furthermore, we found no signal drop in the control condition (physical condition) relative to the ToM conditions, consistent with the Hyper-ToM hypothesis of schizophrenia.

Behavioural impairments in patients with paranoid schizophrenia

As hypothesized the patient group showed lower accuracy and increased reaction times compared to the control group. However, reduced performance was observed in all four conditions, i.e. also the control condition (Ph-C). Brunet *et al.* (2003) using a ToM task similar to ours, reported similar results, i.e. reduced performance in ToM as well as the control task. One could argue that this result reflects unspecific impairments in the patient group. However, an alternative explanation can be provided which interprets these findings as a consequence of the patients exhibiting 'hyper-ToM'. 'Hyper-ToM' leads these patients to attribute intentions and goals to objects, i.e. treating things like persons.

General impairment of neural mechanisms supporting ToM

The task used elicited strong activation in healthy controls in the typical ToM network as in previous studies (Walter *et al.*, 2004). Although patients also showed activation in parts of this network for CInt, the group comparison revealed significantly less activation in the three main ToM regions, namely the MPFC as well the bilateral TPJ (compare Figure 2, Table 2).

As reduced overall performance might influence neuroimaging results we performed three types of analyses as outlined in the methods sections, incorporating either all trials (Model 1), correct trials only (Model 2) or exactly the same number of correct trials in both groups (Model 3). The results of these analyses did not change our main findings. Therefore, we argue that it is justified to assume that the group differences for different conditions we describe in the next section are explained by the specific type of process induced rather than by task difficulty.

Our results are consistent with findings by Brunet *et al.* (2003) using a similar cartoon paradigm who also found less activation in the right MPFC in patients with schizophrenia. However, as we will discuss below, the study by Brunet *et al.* (2003) mixed the types of intentions which might explain why they could not demonstrate hypoactivation of the TPJ. Our data also fit with the findings of Lee *et al.* (2006) who found less activation of the left MPFC in patients with schizophrenia in the acute phase compared to activations after remission. Similar findings have been described by Marjoram *et al.* (2006) comparing high risk subjects with and without psychotic symptoms. In a very recent study by Brüne *et al.* (2008) however, increased activation of dorsomedial areas, left TPJ and right temporal cortex were found in patients with schizophrenia compared to controls. In that particular study patients with passivity symptoms were investigated which in general are least impaired in ToM tasks. Furthermore, the task used by these authors did not show a robust activation of the ToM network in controls, in particular healthy controls showed no activation of the medial prefrontal cortex, perhaps due to the fact that activation and control task were very similar.

Neural dysfunction of the ToM network is modulated by type of intention

One strength of our paradigm is that it allows to investigate group differences for different types of intentions. For the private intention condition (PInt), representing the most simple ToM condition, we found no group differences. Therefore, on the neural level, patients with schizophrenia present no neural dysfunction for this type of ToM task. Instead, during PSInt, significant group differences in the right TPJ and the MPFC were revealed. Although both private intentions (PInt) and prospective intentions (PSInt) share a common element, namely, a single agent acting in isolation, only PSInt requires the representation of a social goal. Also in CInt, the second and more complex social intention involving two people interacting, there was an additional group difference in activation in the left TPJ (together with the right TPJ and the MPFC).

How can we understand these group differences which are restricted to social intentions only? Our results show that as soon as social interaction is involved (present or foreseen) neural differences in activation become apparent. This is true for the most basic structure of the ToM network, namely the right TPJ (Saxe and Kanwisher, 2003; Saxe and Wexler, 2005). Significant group differences are also found in the

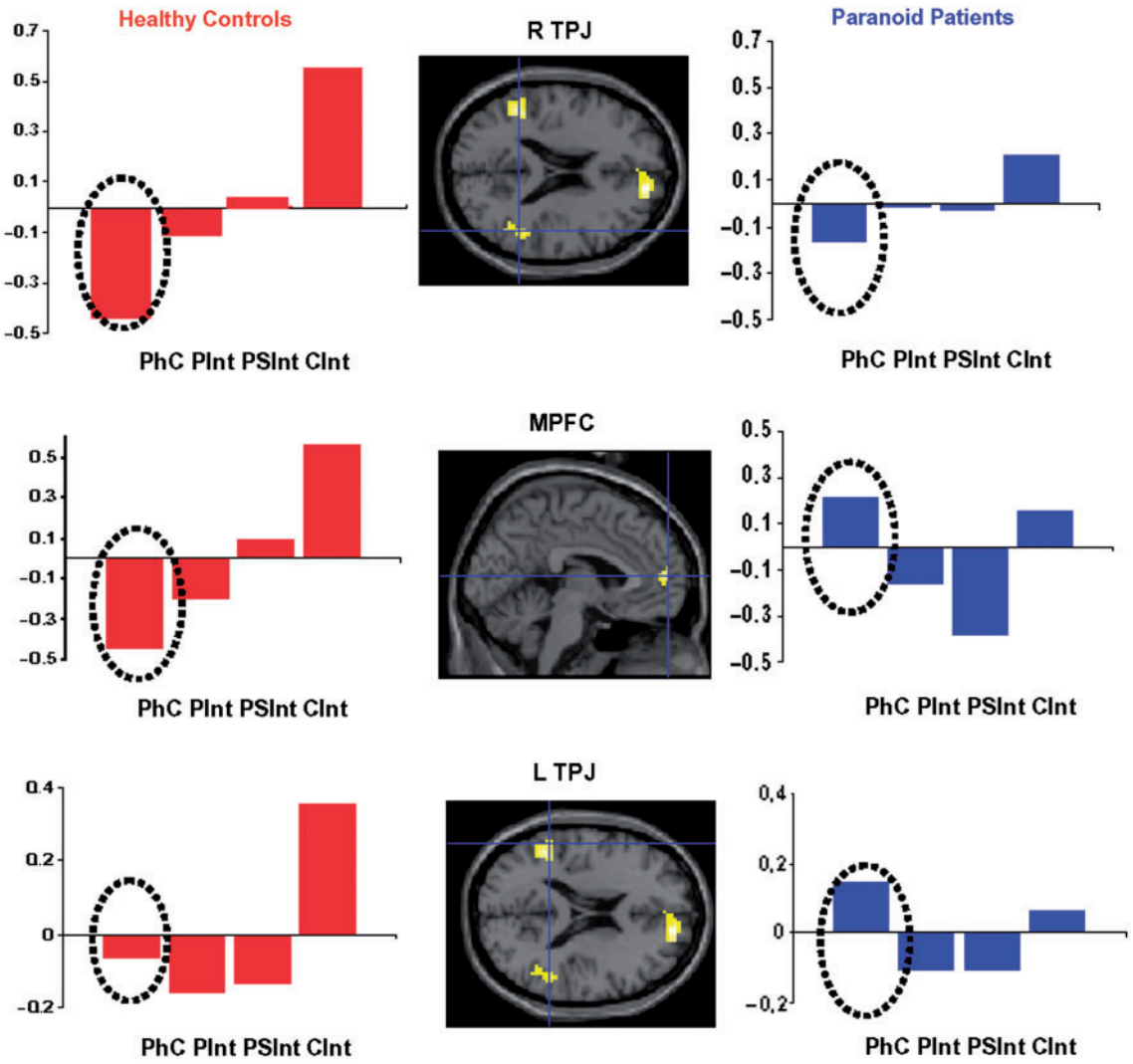

Fig. 5 Mean activation effects (estimated beta parameters, 95% confidence interval) of the contrast CInt vs PhC for right TPJ, MPFC and left TPJ. The activation effects were extracted from the second level between-group ANOVA, $P < 0.001$ uncorrected, $P < 0.05$ at the cluster level. Red: Healthy control subjects; Blue: Paranoid patients. Dashed circles indicate the beta parameters for the control condition (Ph-C).

medial prefrontal cortex (MPFC). It has been argued that the MPFC serves the purpose of decoupling mental states from their environment (Leslie, 1987; Brüne, 2005; Brunet-Gouet and Decety, 2006). This function helps subjects to distinguish clearly between what is happening in the outer world and the inner world. That is especially important in understanding social intentions in which observing subjects have to distinguish between intentions of others and their own as in PSInt and CInt. Patients with schizophrenia show an aberrant pattern of activation for social intentions probably because they are not able of decoupling and have problems distinguishing between intentions of others interacting and their own, resulting in misattributions.

In the CInt condition group differences were most pronounced and incorporated additionally the left TPJ (Figure 3). The left TPJ has been shown to be specifically activated for communicative intentions (Walter *et al.*, 2004; Ciaramidaro *et al.*, 2007). We propose that this specific activation is based on the role of the left hemisphere in language processing. It has been pointed out that ToM deficits in schizophrenia might explain some of the communication problems that patients with schizophrenia have (Frith, 2004). Paranoid schizophrenic patients make significantly more mistakes on tasks that involve inferring the beliefs and intentions of the speakers (Tenyi *et al.*, 2002). In line with this reasoning, Langdon *et al.* (2002), using a picture-sequence task, have demonstrated pragmatic deficits of expressive language and pragmatic deficits of comprehension. Brüne *et al.* (2005) propose that overattribution of intentions is related to incorrect use of pragmatic rules in their use of language.

Hyper-ToM for physical events?

Interestingly, our data reveal further information related to the control condition, i.e. representing physical causality

without any intentions (Ph-C). Looking at the beta parameters in Figure 4, it is obvious that the lack of activation in the MPFC for the contrast CInt vs. Ph-C is not only due to decreased activation in this region in the CInt condition, but also to relatively increased beta values in the Ph-C condition. In accordance with the above-mentioned hypothesis that paranoid patients may have a hyperactive intention detector, we can explain our results tentatively as follows: Paranoid patients do not deactivate their intention detector when they are solving stories involving physical causality but these patients are always in an 'online' modus of ToM. This would also be the case in contexts without intentional agents, where no ToM is required. Blakemore *et al.* (2003) reported that patients with delusions of persecution attributed intentional behaviour to moving shapes in conditions where controls saw no intentionality. These authors propose that patients with schizophrenia perceive agency where others see none. The same process took place when our patient group observed Ph-C stories. An exaggerated sense of agency seems to characterize patients with delusions of persecution, and this tendency to perceive agency where there is none may be a more general feature of schizophrenia (Frith, 2005). This could be the reason why our patients make as many errors in the Ph-C conditions as in the ToM conditions: One may speculate that these patients do not properly recognise the difference between stories involving an intentional agent from stories without an agent: Moving objects are processed as possessing intentional agency.

Also the beta parameters for the left TPJ (Figure 5) showed a similar trend as in the MPFC: We observed an increasing activation in the CInt condition for the patients group (like in the healthy group) but also in the Ph-C condition (in the opposite direction as the healthy group). We speculate that the relative increase in the left TPJ and MPFC for the Ph-C condition suggests that the patient group attributes communicative intentions to the objects present in this scenario (all the Ph-C comic strips contain two or more objects involved in physical causality). Therefore, an important conclusion from our data is that the missing differential activation might result partly from a dysfunction during physical causality attribution and not only from a dysfunction during the attribution of social intentions.

Our proposal is that the ToM-module in patients with paranoid schizophrenia might malfunction because it is overactive from the start and thus is not well suited to distinguish properly between mental and physical states. Hence, we agree with the idea of 'hyper ToM' as proposed by Abu-Akel and Bailey (2000): 'An attitude to associate with quantitative overgeneration of hypotheses or overattribution of mental states' also when ToM is not demanded.

Our study has several limitations. First, we studied only one subgroup of patients with schizophrenia, namely paranoid schizophrenia. The relatively small number of patients might also explain why we were not able to demonstrate correlations with psychopathology. It would be helpful to investigate different subtypes of schizophrenia in future studies, in particular disorganized patients. Secondly, we did not include an exhaustive behavioural battery of ToM tasks which might allow interpreting our results in relation to behavioural findings. Thirdly, we have incorporated only patients that were medicated so that we cannot definitely exclude medication effects. However, as our findings were not general but specific to certain conditions within the patient group our results are unlikely to be explained only by medication.

CONCLUSION

In summary, we have demonstrated that different parts of the ToM network are differentially affected in understanding the social domain in relation to the type of intention involved. Whereas we found no group differences for private intentions, group differences emerged for social intentions, namely reduced differential activation for prospective intentions in the right TPJ and the MPFC and, for communicative intentions additionally in the left TPJ. We provide evidence that the dysfunction in the intentional network is partially mediated by an intention detector which became hyperactive in the paranoid interpretation of the physical world. Our results clearly demonstrate that findings of dysfunctional activation within the ToM network can only be interpreted properly if the type of mental state or intention which is used to solve the task is taken into account. Furthermore, our results provide evidence that the dispute about the true mentalizing structure within the brain might be a misnomer: Different structures might be relevant for different types of intentional states. Regarding the social and communicative aspects of ToM tasks we have provided evidence that the MPFC and the left TPJ play the most important role. This might stimulate research in different ways. First, it would be valuable to use private and communicative intentions in order to study subtypes of schizophrenia e.g. disorganized versus negative symptoms versus positive symptoms. Second, our results point to the necessity to look also at the control condition used in ToM tasks, especially in patients with positive symptoms which might have a hyperactive intention detector already for physical events.

REFERENCES

Abu-Akel, A., Bailey, A.L. (2000). The possibility of different forms of theory of mind. *Psychological Medicine*, *30*, 735–738.

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Review of Neuroscience*, *7*, 268–277.

Andreasen, N.C., Carpenter, W.T., Jr, Kane, J.M., Lasser, R.A., Marder, S.R., Weinberger, D.R. (2005). Remission in schizophrenia: Proposed criteria and rationale for consensus. *American Journal of Psychiatry*, *162*, 441–449.

Apperly, I.A., Samson, D., Chiavarino, C., Humphreys, G.W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: Neuropsychological evidence from a false-belief task with reduced

language and executive demands. *Journal of Cognitive Neuroscience*, *16*, 1773–1784.

Baron-Cohen S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*: Cambridge, MA: MIT Press.

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. *Cerebral Cortex*, *13*, 837–844.

Brothers, L. (1990). The social brain: A project for integrating primate behaviour and neurophysiology in a new domain. *Concepts Neuroscience*, *1*, 27–51.

Brüne, M. (2001). Social cognition and psychopathology in an evolutionary perspective. Current status and proposals for research. *Psychopathology*, *34*, 85–94.

Brüne, M. (2005). ''Theory of mind'' in schizophrenia: A review of the literature. *Schizophrenia Bulletin*, *31*, 21–42.

Brüne, M., Lissek, S., Fuchs, N., et al. (2008). An fMRI study of theory of mind in schizophrenic patients with ''passivity'' symptoms. *Neuropsychologia*.

Brunet-Gouet, E., Decety, J. (2006). Social brain dysfunctions in schizophrenia: A review of neuroimaging studies. *Psychiatry Research*, *148*, 75–92.

Brunet, E., Sarfati, Y., Hardy-Bayle, M.C., Decety, J. (2003). Abnormalities of brain function during a nonverbal theory of mind task in schizophrenia. *Neuropsychologia*, *41*, 1574–1582.

Ciaramidaro, A., Adenzato, M., Enrici, I., et al. (2007). The intentional network: How the brain reads varieties of intentions. *Neuropsychologia*, *45*, 3105–3113.

Corcoran, R., Cahill, C., Frith, C.D. (1997). The appreciation of visual jokes in people with schizophrenia: A study of 'mentalizing' ability. *Schizophrenia Research*, *24*, 319–327.

Dunbar, R.I.M. (1998). The social brain hypothesis. Evolutionary anthropology. *Evolutionary Anthropology*, *6*, 178–190.

Duvernoy, H.M. (1999). *The Human Brain*. Springer, Wien.

Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C. (1995). Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of cluster-size threshold. *Magnetic Resonance Medicine*, *33*, 636–647.

Frith C.D. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Lawrence Erlbaum Associates Ltd., Hove.

Frith, C.D. (2005). The self in action: Lessons from delusions of control. *Conscious Cognitive*, *14*, 752–770.

Frith, C.D. (2004). Schizophrenia and theory of mind. *Psychology Medicine*, *34*, 385–389.

Frith, C.D., Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*, 531–534.

Gallagher, H.L., Frith, C.D. (2003). Functional imaging of 'theory of mind'. *Trends in Cognitive Sciences*, *7*, 77–83.

Gregory, C., Lough, S., Stone, V., et al. (2002). Theory of mind in patients with frontal variant frontotemporal dementia and Alzheimer's disease: Theoretical and practical implications. *Brain*, *125*, 752–764.

Hynes, C.A., Baird, A.A., Grafton, S.T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia*, *44*, 374–383.

Kay, S.R., Opler, L.A., Lindenmayer, J.P. (1989). The Positive and Negative Syndrome Scale (PANSS): Rationale and standardisation. *British Journal of Psychiatry Supplement*, 59–67.

Langdon, R., Coltheart, M., Ward, P.B., Catts, S.V. (2002). Disturbed communication in schizophrenia: The role of poor pragmatics and poor mind-reading. *Psychology. Medicine*, *32*, 1273–1284.

Lee, K.H., Brown, W.H., Egleston, P.N., et al. (2006). A functional magnetic resonance imaging study of social cognition in schizophrenia during an acute episode and after recovery. *American Journal of Psychiatry*, *163*, 1926–1933.

Leslie, A.M. (1987). Pretence and representation: The origins of 'theory of mind. *Psychological Review*, *94*, 412–426.

Marjoram, D., Job, D.E., Whalley, H.C., et al. (2006). A visual joke fMRI investigation into Theory of Mind and enhanced risk of schizophrenia. *Neuroimage*, *31*, 1850–1858.

Overall, J.E., Gorham, D.R. (1976). In: Guy, W., editor, *BPRS. Brief Psychiatric Rating Scale. ECDEU, Assessment Manual for Psychopharmacology*, Rockville: National Insitute of Mental Health, pp. 157–169.

Pickup, G.J., Frith, C.D. (2001). Theory of mind impairments in schizophrenia: Symptomatology, severity and specificity. *Psychology Medicine*, *31*, 207–220.

Premack, D., Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behaviour & Brain Science*, *1*, 515–526.

Rowe, J.B., Passingham, R.E. (2001). Working memory for location and time: Activity in prefrontal area 46 relates to selection rather than maintenance in memory. *Neuroimage*, *14*, 77–86.

Russell, T.A., Rubia, K., Bullmore, E.T., et al. (2000). Exploring the social brain in schizophrenia: Left prefrontal underactivation during mental state attribution. *American Journal of Psychiatry*, *157*, 2040–2042.

Saxe, R. (2006). Uniquely human social cognition. *Current Opinion Neurobiology*, *16*, 235–239.

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.

Saxe, R., Wexler, A. (2005). Making sense of another mind: The role of the right temporo-parietal junction. *Neuropsychologia*, *43*, 1391–1399.

Sprong, M., Schothorst, P., Vos, E., Hox, J., van Engeland, H. (2007). Theory of mind in schizophrenia: Meta-analysis. *British Journal of Psychiatry*, *191*, 5–13.

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., Jr, Alexander, M.P. (2001). The frontal lobes are necessary for 'theory of mind'. *Brain*, *124*, 279–286.

Talairach J., Tournoux P. (1988). *Coplanar Stereotaxic Atlas of the Human Brain*. Thieme, Stuttgart.

Tenyi, T., Herold, R., Szili, I.M., Trixler, M. (2002). Schizophrenics show a failure in the decoding of violations of conversational implicatures. *Psychopathology*, *35*, 25–27.

Walter, H., Adenzato, M., Ciaramidaro, A., Enrici, I., Pia, L., Bara, B.G. (2004). Understanding intentions in social interaction: The role of the anterior paracingulate cortex. *Journal of Cognitive Neuroscience*, *16*, 1854–1863.