



ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage xx (2007) xxx–xxx

Neural correlates of true and false belief reasoning

Monika Sommer,^{a,*} Katrin Döhnelt,^a Beate Sodian,^b Jörg Meinhardt,^b
Claudia Thoermer,^b and Göran Hajak^a

^aDepartment of Psychiatry and Psychotherapy, University of Regensburg, Universitätsstraße 84, D-93053 Regensburg, Germany

^bDepartment of Psychology, Ludwig-Maximilians University, Munich, Germany

Received 24 August 2006; revised 20 December 2006; accepted 14 January 2007

Belief reasoning plays a central role in making inferences about other people's mental states. The ability to reason about false beliefs is considered as a critical test for having a Theory of Mind (ToM). There is some controversy as to whether it is the medial prefrontal cortex (MPFC) or the temporo-parietal junction (TPJ) that is centrally involved in belief reasoning. According to developmental studies of belief reasoning we conducted an fMRI experiment with a carefully controlled paradigm (Sally Anne scenario). We compared false belief reasoning with true belief reasoning in parallel tasks, using a series of cartoon stories depicting transfer of an object unbeknownst to the protagonist (false belief) or with the protagonist witnessing (true belief). The false belief versus true belief contrast revealed activation of the dorsal part of the anterior cingulate cortex (dACC), the right lateral rostral prefrontal cortex and the right TPJ associated with false belief. We suggest that the activation of the dACC and the lateral PFC might be associated with action monitoring and stimulus-independent cognitive processing whereas the activation of the TPJ might be related to the computation of mental representations that create perspective differences, such as a person's false belief that contrasts with reality and therefore might be centrally involved in the decoupling mechanism. Additionally we found common patterns of activation for true and false belief reasoning, including inferior parietal and precuneus activation, but we found no activation of the MPFC or the TPJ in general belief reasoning.

© 2007 Elsevier Inc. All rights reserved.

Introduction

The fundamental human ability to attribute mental states to oneself and others, known as Theory of Mind (ToM), has recently become an active area of functional neuroimaging research (Gallagher and Frith, 2003; Saxe et al., 2004; Amodio and Frith, 2006; Frith and Frith, 2006 for reviews).

The ability to represent false beliefs is commonly considered as the critical test for a Theory of Mind since it involves a differentiation between belief and reality. Only when a false belief

is represented independently of the real state of affairs can action be predicted or explained correctly. In contrast, when people hold true beliefs belief-representation is not necessary to predict or explain their actions; such predictions could simply be derived from a representation of reality.

Consequently, neuroimaging studies dealing especially with belief reasoning have commonly used false belief stories or cartoons and contrasted these with sequences of events that do not invite mentalizing (i.e., physical or mechanical reasoning) (Fletcher et al., 1995; Gallagher et al., 2000; Goel and Dolan, 2001; Vogeley et al., 2001; Saxe and Kanwisher, 2003; Grèzes et al., 2004). Using this approach, studies have gathered evidence that the medial prefrontal cortex (MPFC), especially the anterior paracingulate area, and the temporo-parietal junction (TPJ) are involved in belief attribution. Nevertheless the results are inconsistent. Using verbal stories some studies found activation of the anterior cingulate cortex (Fletcher et al., 1995; Vogeley et al., 2001) or the anterior part of the rostral prefrontal cortex (Saxe and Kanwisher, 2003) whereas studies using cartoons or films found activation of the more dorsal part of the rostral prefrontal cortex (Gallagher et al., 2000; Grèzes et al., 2004). Inconsistencies also concern the temporal and parietal areas. In contrast to non-mentalizing conditions false belief tasks activated the temporo-parietal junction (Fletcher et al., 1995; Gallagher et al., 2000; Saxe and Kanwisher, 2003) or the more posterior parts of the superior temporal areas (Grèzes et al., 2004).

One reason for the different activation patterns could be the use of different subtracted non-ToM tasks. Fletcher's "physical" stories (Fletcher et al., 1995) consisted to some extent of true belief stories, which described actions based on the character's true belief and to some extent of physical stories without any kind of belief. Gallagher et al. (2000) used non-ToM cartoons, which might induce mentalizing, because participants may try to understand the cartoonist's intended joke. Saxe and Kanwisher (2003) compared the false belief ToM conditions with two non-mentalizing conditions, mechanical inference and false photographs. A further reason for the different results could be that the experimental belief attribution tasks involved not only belief attribution but also other kinds of attribution processes (e.g., intention attribution, emotion attribution). Verbal stories contain complex plots (Fletcher et al., 1995; Gallagher et al., 2000; Saxe and Kanwisher, 2003) and do

* Corresponding author. Fax: +49 941 9412065.

E-mail address: monika.sommer@medbo.de (M. Sommer).

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

not only involve belief attribution but also intention and emotion attribution. Because of their humorous content, cartoons (Gallagher et al., 2000) may induce pleasant emotions and different attribution processes.

Saxe et al. (2004) argue that a specialized neural substrate for reasoning about beliefs has to meet two criteria: generality and specificity. While previous neuroimaging studies have addressed the specificity criterion, by contrasting false belief reasoning with various non belief-related control conditions, the generality issue has been neglected. The candidate brain regions should not only show increased activity in response to false belief attribution, but should also be activated in reasoning about true beliefs. Some studies had control conditions which involved true belief reasoning with respect to a protagonist's action (Fletcher et al., 1995; Gallagher et al., 2000; Saxe and Kanwisher, 2003), but as mentioned above the false belief and the control tasks generally differed on more than one dimension (e.g., false belief stories and control stories were taken from different content domains). In contrast, developmental studies have contrasted true and false belief reasoning in carefully controlled paradigms (see Wellman et al., 2001; Sodian, 2005 for an overview). To date, no neurocognitive studies of theory of mind reasoning in children have been reported. Comparisons of neural activity in children and adults involved in true and false belief reasoning are needed in order to determine whether there is a specialized system for theory of mind reasoning even in young children and under what circumstances this development takes place. As a first step towards such developmental comparisons we need to study true and false belief reasoning in adults in a task that exclusively involves belief reasoning and that could also be used with children. We therefore used a series of cartoon stories depicting unexpected transfer events (an object is transferred from location A to B unbeknownst to the protagonist (false belief condition) or with the protagonist witnessing the transfer (true belief condition)). The comparison of neural activation associated with false and true belief events allows for a critical test of the association of brain regions involved in true and false belief reasoning. Furthermore, the paradigm also tests for the specificity of false belief over true belief reasoning. Cognitive accounts of false belief reasoning have emphasized the decoupling of mentality and reality involved in false belief attribution. Since previous neuroimaging studies did not explicitly contrast true and false belief reasoning, we do not know about the brain activation associated specifically with decoupling mental states from reality.

Therefore, the aim of the present study is to explore the neural correlates involved in belief attribution in general and specifically in decoupling mental states from reality as it is necessary in false belief attribution.

Methods

Subjects

Sixteen subjects (eight men and eight women; range 23–37 years, mean 26) with no neurological or psychiatric history participated in the imaging study. All gave informed consent according to the guidelines of the local Ethic Committee.

Stimuli

We presented nonverbal cartoon stories depicting a person acting on the basis of correct (true belief) or incorrect (false belief)

representations of reality (Fig. 1). All stories consist of 7 pictures and tell a story according to the “Sally Anne Scenario” (Baron-Cohen et al., 1985). The first four pictures show two children (e.g., Betty and Nick) and two boxes (e.g., a basket and a bag) in a room (picture 1). One child (e.g., Betty) puts an object into one box (e.g., a ball into the bag; picture 2). Then Betty goes out of the room (picture 3). The other child (Nick) takes the ball out of the bag (picture 4). These four pictures are the same in the false and the true belief conditions. Then the stories continue slightly differently. In the true belief condition Betty comes back into the room (picture 5) and observes that Nick puts the ball into the basket (picture 6). In the false belief condition Nick puts the ball into the basket (picture 5) and then Betty comes back into the room (picture 6). The response picture (picture 7) at the end of each story was marked by a red frame and was the same for the true and the false belief story: Betty looks for the ball. In 50% of the trials based on her belief Betty looks into the correct box (expected) and in 50% of the trials based on her belief she looks into the false box (unexpected). The subjects were required to decide whether on the basis of Betty's belief the end of the story was expected or unexpected by making key press. Responses were recorded using two buttons of a five-button fMRI compatible response pad. Subjects used the index and the middle finger of the right hand for response. The allocation of the finger to the answer (expected or unexpected) was counterbalanced.

Each fMRI session consisted of two runs. In each run 20 true belief stories and 20 false belief stories were randomly presented with an intertrial interval varying randomly between 5.6 and 13.8 s. Altogether every subject saw 40 true and 40 false belief stories.

Each picture was presented for 1400 ms with the exception of the last picture, which was presented for 2000 ms. The stories were divided in three sections according to their content and the underlying information processes (Fig. 1). The first section (story) contained pictures 1–4, the second (belief) pictures 5–6 and the last (response) picture 7. The different sections were divided by intertrial intervals varying randomly between 5.6 and 13.8 s.

The software package Presentation (Neurobehavioural Systems Inc., <http://www.neuro-bs.com>) was used to present stimuli and to record responses. The stories were viewed by projection onto a mirror mounted onto the head coil in the scanner.

fMRI

Imaging was performed using a 1.5 T scanner (Siemens Sonata, Erlangen, Germany). The functional images sensitive to blood oxygenation level-dependent (BOLD) contrasts were acquired by T2*-weighted echo planar (EPI) images (TR=2.82 s, TE=60 ms, $\alpha=90^\circ$ in plane matrix 64×64 ; FoV=192 mm). The images consisted of 30 contiguous axial slices with 3 mm thickness and $3 \text{ mm} \times 3 \text{ mm}$ in plane resolution. During each run 525 volumes were continuously acquired. The first four volumes were automatically discarded to allow for T1 equilibration effects. High-resolution structural T1-weighted MPAGE images (TR=1.9 s, TE=3.93 ms, TI=1100 ms, voxel size $1 \times 1 \times 1 \text{ mm}$, 176 axial slices) were also acquired on all subjects.

Statistical analyses

All fMRI data were processed using SPM2 software package (<http://www.fil.ion.ucl.ac.uk>). For each subject, the 1042 image volumes of the two runs were corrected for acquisition timing

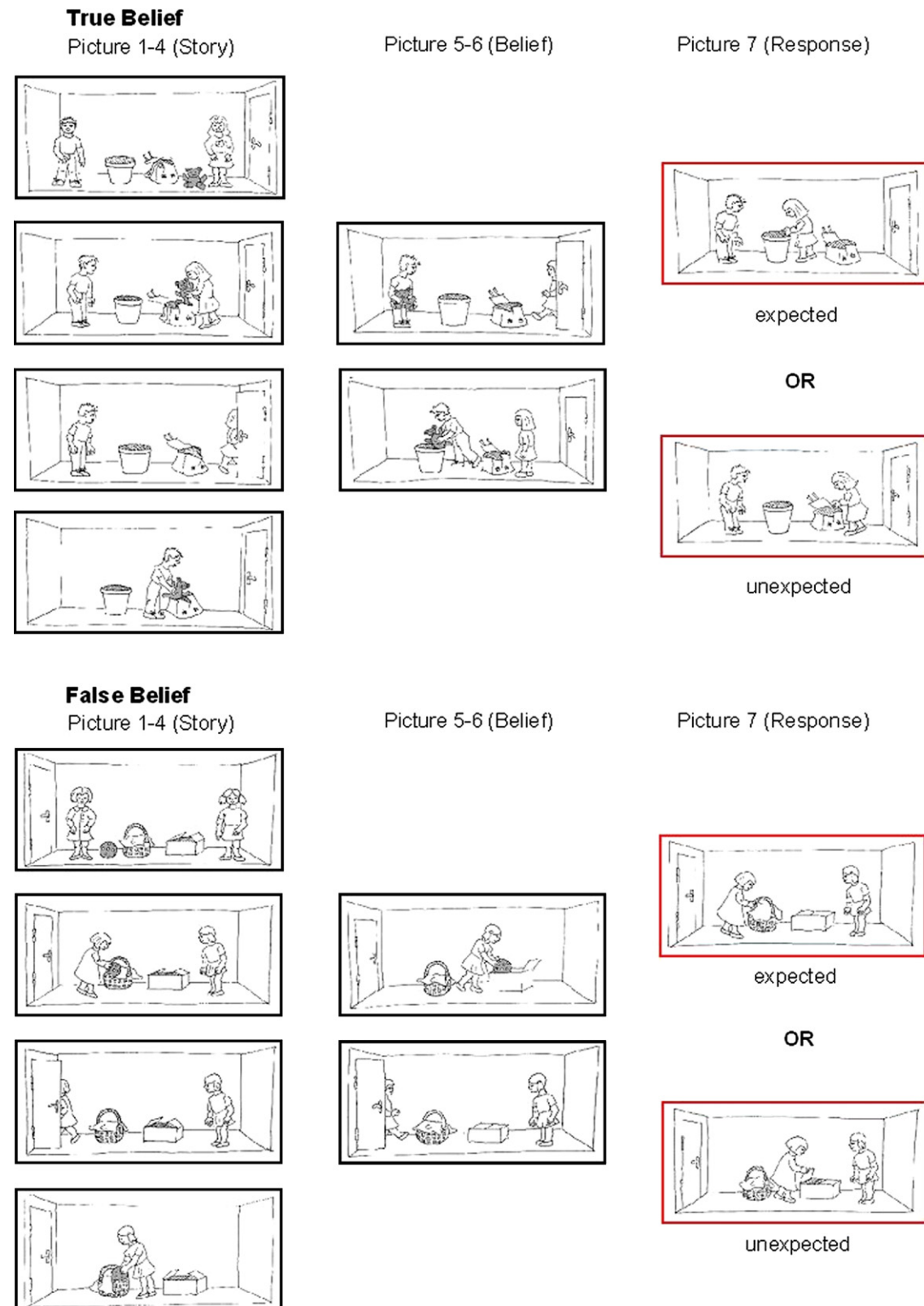


Fig. 1. Example of a true belief story (top) and a false belief story (bottom).

using the middle slice as reference (Henson et al., 1999) and were realigned to the first volume by rigid body transformation to correct for subjects motion. Images were stereotactically normal-

ized to the Montreal Neurological Institute (MNI) reference brain using the mean of the 1042 images and spatially smoothed with a 8 mm full-width half-maximum Gaussian filter.

A first-level fixed effects analysis was computed for each subject using the general linear model with hemodynamic response function modeled by a box car waveform convolved with the hemodynamic response function and its temporal and dispersion derivatives to account for any temporal shifts in the response of the stimuli (Friston et al., 1998). Also included for each session were six covariates to capture residual movement-related artefacts, and a single covariate representing the mean (constant) over scans. The data were high-pass filtered with a frequency cut-off at 128 s. We used the story sections (pictures 1–4) as baseline condition. These pictures have the same visual complexity as the belief pictures (pictures 5–6), but they contain no belief attribution. Four contrasts of interest were computed for every subject: true belief (“true belief” versus “story”), false belief (“false belief” versus “story”) and belief (“false belief” versus “true belief” and “true belief” versus “false belief”). The response picture was only used for measuring behavioral data and was not involved in the fMRI analysis.

Single-subjects’ first-level contrasts were introduced in second-level random-effect analyses to allow for population inference. Main effects were computed using one-sample *t*-test. The resulting set of voxel values for each contrast constituted an SPM map. The maps’ threshold was set at $p < 0.05$ and corrected with false discovery rate (FDR) for multiple comparisons.

Overlaps between true and false belief were determined by computing a conjunction analyses of the true belief (true belief story) and false belief (false belief story) contrasts, using procedures provided by Thomas Nichols (<http://www.sph.umich.edu/~nichols/Conj>). This method provides a valid test for a logical “and” of effects as it allows the rejection of the null hypothesis that one or more of the comparisons have not activated even under dependence between the tests (Nichols et al., 2005). The conjunction analysis was performed at the second level, that is, in contrast to images obtained from the single subject analysis, and was therefore a random-effect analysis. The map’s threshold was set at $p < 0.05$ and corrected with FDR for multiple comparisons. All maps were overlaid on the MNI template.

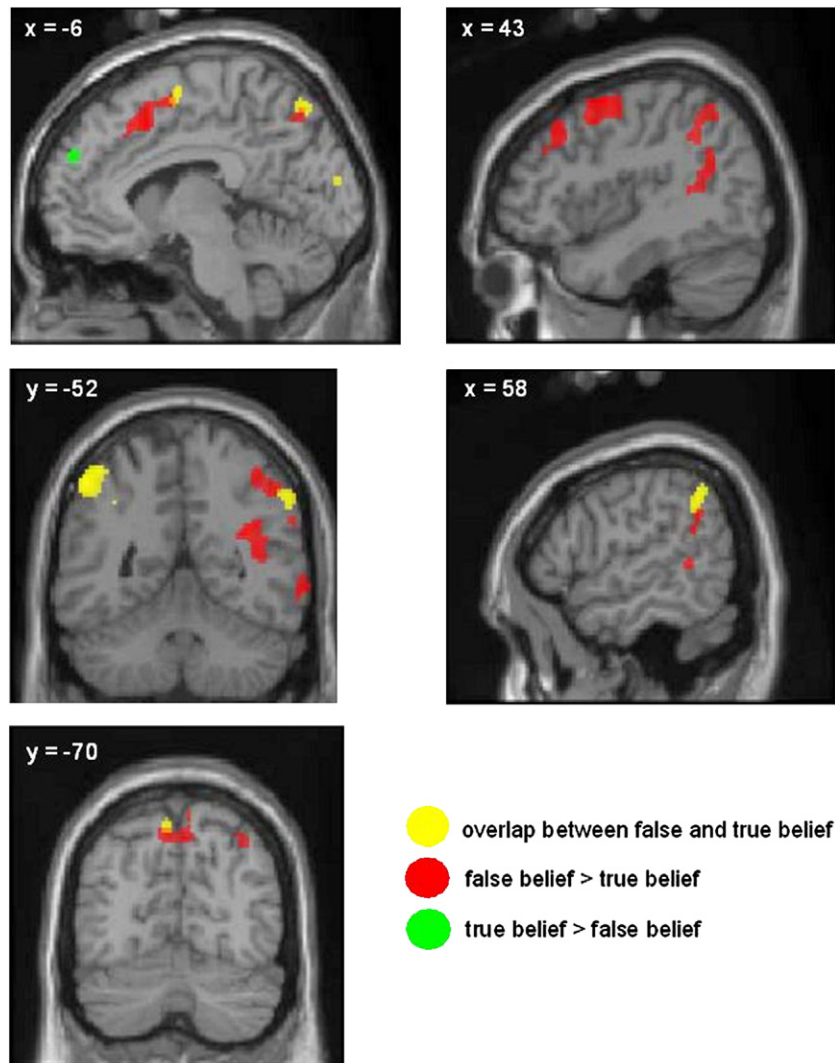


Fig. 2. Shared representational network for false and true belief (yellow) and areas specifically activated for false belief (red) and true belief (green). The group data are superimposed on the MNI brain template.

Results

Behavioral measures

Response accuracy (in percent) and reaction time (in ms) were analyzed by two separate repeated measures analyses of variance (ANOVA) with the two factors “belief” (true belief/false belief) and “expectation” (expected/unexpected). For response accuracy there were no effects for “belief” ($F(1,15)=0.56$, n.s.), “expectation” ($F(1,15)=1.13$, n.s.) or the “belief” \times “expectation” interaction ($F(1,15)=0.89$, n.s.). Participants gave 97% (SD=0.04%) correct answers in the true belief condition and 96% (SD=0.04%) in the false belief condition. For reaction time there were no effects for “belief” ($F(1,15)=0.15$, n.s.) or the “belief” \times “expectation” interaction ($F(1,15)=0.01$, n.s.). But participants reacted significantly faster for the expected end of the story ($M=800.4$ ms, SD=129.6 ms) than for the unexpected end ($M=894.8$ ms, SD=178.1 ms; $F(1,15)=11.22$, $p<0.01$). On the basis of the behavioral data we considered that for adult subjects the true and false belief conditions are equally difficult.

Functional imaging data

The shared representation network for true and false belief reasoning revealed by the conjunction analysis is shown in Fig. 2 (yellow color) and local maxima of activated foci are given in Table 1. Significant overlaps were located in the left superior, middle and inferior frontal gyrus, the bilateral inferior parietal gyrus, the precuneus and the cuneus.

Direct comparisons of the two conditions of interest highlighted significant activations for the false belief minus true belief contrast

in the medial dorsal anterior cingulate cortex, the right middle frontal gyrus, the right dorsolateral prefrontal cortex, the right lateral rostral prefrontal cortex, the right temporo-parietal junction, the right middle temporal gyrus, the right inferior parietal gyrus and the precuneus (Fig. 2, red color, Table 1). The contrast true belief minus false belief showed significantly more activation in the medial rostral prefrontal cortex (Fig. 2, green color, Table 1).

Discussion

Although the neural correlates of ToM have recently become an active area of functional neuroimaging research (Frith and Frith, 2003; Gallagher and Frith, 2003; Saxe et al., 2004; Saxe, 2006) only few studies have dealt specifically with belief attribution. Following behavioral studies in Developmental Psychology we compared false belief reasoning with true belief reasoning in parallel tasks. False belief reasoning in contrast to true belief reasoning requires the representation of a person’s mental state independently of reality, i.e., a decoupling of mental representation and reality.

First, our results showed that false belief in contrast to true belief induced more activation in the frontal cortex, especially the dorsal part of the anterior cingulate cortex (dorsal ACC), the right middle frontal gyrus, the right dorsolateral prefrontal cortex and the right lateral rostral prefrontal cortex.

The dorsal ACC has been associated with several different cognitive tasks (for reviews see Bush et al., 2000; Steele and Lawrie, 2004). According to Amodio and Frith (2006) the dorsal part of the anterior cingulate cortex can be seen as a part of the posterior region of the rostral medial frontal cortex (prMFC), a region that is involved in representing and continuously updating possible future actions in order to regulate behavior. Although in true and false belief conditions the transfer of an object takes place, in the false belief condition the transfer has a greater impact on participant’s prediction of the child’s future behavior. In the false belief condition monitoring the transfer results in the decoupling of the child’s mental representation from the real state of affairs. The child’s future action will be caused by the mental representation about the object’s location and not by the real location. Therefore it is possible that the found activation of the dorsal ACC for false belief compared to true belief might represent processes associated with action monitoring resulting in the decoupling of the participant’s knowledge about reality and the child’s mental representation.

The representation of the child’s visual access during the object’s removal is essential to understand whether the child’s belief, with respect to reality, is false or true. In addition to triggering processes of action monitoring, false belief reasoning in such tasks can be seen as a case of stimulus-independent task processing. Corresponding to the found activation of the dorsal ACC (−8, 16, 46) in response to false belief, Gilbert et al. (2006a) found increased activation of the bilateral ACC (−8, 14, 46) in response to a stimulus-independent condition in contrast to a stimulus-oriented condition. Interpreting our results in terms of stimulus-independent versus stimulus-oriented processing, the central process in the false belief condition is the mental representation of the child’s belief where the object is independent of the real location of the object. Therefore, the activation of the dorsal ACC may be correlated with both the monitoring of the action (Amodio and Frith, 2006) and the necessary stimulus-independent processing of the child’s mental representation (Gilbert et al., 2006a).

The interpretation that a stimulus-independent cognitive process is involved during false belief reasoning is additionally

Table 1
Regions of increased brain activity associated with belief attribution

Brain region	Center MNI coordinates			Z-score	Voxels (n)
	x	y	z		
<i>Conjunction of true and false belief</i>					
Left superior frontal gyrus (BA 6)	−22	8	66	5.55	157
Left middle frontal gyrus (BA 46)	−42	32	28	5.44	47
Left inferior frontal gyrus (BA 47)	−52	14	0	5.48	43
Left inferior parietal lobe (BA 40)	−52	−50	48	6.18	334
Right inferior parietal lobe (BA 40)	56	−50	42	5.94	111
Precuneus (BA 7)	−6	−72	56	5.46	37
Cuneus (BA 18)	−2	−90	14	5.25	49
<i>False belief > True belief</i>					
Dorsal anterior cingulate cortex (BA 32)	−8	16	46	4.62	358
Right middle frontal gyrus (BA 6)	42	4	58	4.56	373
Right dorsolateral prefrontal cortex (BA 9)	46	28	40	4.05	202
Right lateral rostral prefrontal cortex (BA 10)	34	64	6	3.74	153
Right temporo-parietal junction (BA 39)	34	−54	24	4.29	311
Right middle temporal gyrus (BA 21)	64	−54	−4	4.14	175
Right inferior parietal gyrus (BA 7)	38	−64	46	3.97	530
Precuneus (BA 7)	4	−74	52	3.77	157
<i>True Belief > False belief</i>					
Medial rostral prefrontal cortex (BA 10)	0	56	26	3.81	177

Coordinates refer to the Montreal Neurological Institute (MNI) reference brain. Brodmann areas (BA) are approximate.

supported by the increased activation of the right lateral rostral prefrontal cortex in response to false belief. Gilbert et al. (2005) suggest that the lateral rostral PFC is involved in switching attention between stimulus-oriented and stimulus-independent thought processes. The observed activation of the right lateral PFC in the false belief condition in contrast to the true belief condition may represent this switching of attention between the reality based representation of the stimulus and the reality independent mental representation of the looking child's belief.

Further support for the interpretation of belief reasoning in terms of stimulus-dependent and stimulus-independent mental processes comes from the contrast “true belief versus false belief”. In the present study true belief in contrast to false belief induced activation of the rostral medial prefrontal cortex (MPFC; BA 10). Gilbert et al. (2005) showed that MPFC activation was greater in phases requiring stimulus-oriented thought than in phases requiring stimulus-independent thought. These results are consistent with the present finding of increased activation during true belief attribution. In the true belief condition, the child's mental representation of the object's location and the representation of reality in the participant processing the stimulus were the same. The participant did not have to process the mental state of the child independently of reality. This allows stimulus-oriented processing resulting in increased activation of the MPFC. In contrast, false belief requires the stimulus-independent decoupling of mental state and reality, resulting in increased activation of the lateral PFC.

Second, contrasting false belief versus true belief we found increased activation of the right temporo-parietal junction (TPJ-R). Studies investigating false belief consistently revealed activation of the TPJ-R (Frith and Frith, 2003; Saxe and Kanwisher, 2003; Perner et al., 2006). Saxe and Kanwisher (2003) suggest that the TPJ responds selectively to mental states. They showed that the region activated more when subjects read stories about false beliefs than during reading false photograph stories and that the area showed more activation in response to stories about a character's mental state than to stories describing people in physical detail. In the present study the TPJ-R was not involved in general belief reasoning, but specifically activated during false belief. Therefore our results did not support the assumption by Saxe and Kanwisher (2003) and Saxe (2006) that the TPJ-R is generally associated with mentalizing. Investigating different kinds of representations (false beliefs, false photos, false signs) Perner et al. (2006) suggest that the TPJ-R is involved in computing mental states that create a perspective difference, such as a person's false belief that contrasts with the state of reality. The results of the present study support this interpretation and emphasize the role of the TPJ-R in the decoupling mechanism.

Third, the contrast “true belief minus false belief” revealed activation of the medial rostral PFC. Although previous neuroimaging studies did not directly compare true belief to false belief, a large number of studies investigating the neural correlates of mentalizing consistently found activation in this region (Frith and Frith, 2003; Amodio and Frith, 2006; Frith and Frith, 2006). These studies included different mentalizing conditions, like intention attribution to animated shapes (Castelli et al., 2000), depicted people (Walter et al., 2004) or imaginary game partners (McCabe et al., 2001; Gallagher et al., 2002), inferring the beliefs of someone who lived in prehistoric times (Goel et al., 1995), or judgments about psychological states (Mitchell et al., 2005). Gilbert et al. (2006b) showed that these mentalizing studies are associated with activations relatively caudal in the rostral MPFC

(mean *y*-coordinate 53.4) whereas cognitive multitask studies are associated with relatively rostral activations (mean *y*-coordinate 57.1). In the present study true belief reasoning induced activation in this more rostral part of the rostral MPFC (*y*-coordinate=56). This activation may be associated rather with cognitive processes than mentalizing.

The main purpose of our study was to investigate common patterns of activation during true and false belief reasoning. Although Fletcher et al.'s physical stories can be seen as a true belief condition, to date a comparison of false and true belief reasoning in parallel tasks with the same physical complexity and a very similar content has not been done in neuroimaging studies. Following Saxe et al. (2004) who argued that a brain region involved in belief attribution should show increased activity in response to any stimuli that invite belief attribution, we used a conjunction analysis to identify overlaps between the activations determined by true and false belief attribution. This analysis revealed increased activation in several brain areas including left superior, middle and inferior frontal gyrus, bilateral inferior parietal lobe, precuneus and cuneus. Contrary to expectations, the conjunction analysis revealed no common patterns of activation for the true and false belief conditions in brain areas assumed to play a central role in mentalizing: the relatively caudal part of the anterior rostral medial frontal cortex or the TPJ.

The missing activation of the caudal parts of the anterior medial rostral PFC may be associated with our kind of belief task. In a meta-analysis of functional neuroimaging studies Gilbert et al. (2006b) investigated functional specialization within the rostral MPFC. According to their analysis, activation in the relatively caudal medial rostral PFC is associated with tasks that involve both mentalizing and emotional material. Mentalizing tasks without emotional material did not activate caudal parts of the medial rostral PFC. In contrast to most mentalizing tasks that include emotional material (Gilbert et al. (2006b) found only 2 of 26 mentalizing studies without emotional material), the belief conditions in the present study did not involve emotional aspects. Both the transfer of the object and the search for the object were always the same and had no emotional consequences.

Furthermore the conjunction analysis did not show activation of the TPJ-R indicating that the TPJ-R does not fulfil the generality criterion of Saxe et al. (2004). But our results support the involvement of the TPJ-R in computing mental states that create perspective differences (Aichhorn et al., 2006; Perner et al., 2006).

Finally, both the false belief versus true belief contrast and the conjunction analysis revealed increased activation of the precuneus. Several recent neuroimaging findings suggest a central role for the precuneus in a wide spectrum of highly integrated tasks, including visuo-spatial imagery, episodic memory retrieval and self processing operations (Vogele et al., 2001; Den Ouden et al., 2005; Cavanna and Trimble, 2006). There are extensive connections between precuneus and frontal cortex, especially BA 8, 9 and 46 and also anterior cingulate cortex (Leichnetz, 2001). Vogele and Fink (2003) suggest that the medial parietal cortex is involved in assigning first-person perspective and interpreting an action as being controlled by oneself versus another person. The ability to distinguish one's own perspective from others' is extremely relevant for belief attribution because one must know that the mental representation of other people's minds can be different from one's own. Possibly, the observed increased activation of the precuneus during both belief conditions represents the process of mental state attribution in both belief conditions. Interestingly, the

precuneus is more activated during false than true belief attribution, which might be correlated with the more demanding processes of decoupling mentality and reality.

Whereas false belief in contrast to true belief induced more activation in the right TPJ (BA 39), the conjunction analysis revealed that belief attribution generally activated bilateral inferior parietal cortex (BA 40). Several studies showed that parietal areas play an important role in perspective taking (Aichhorn et al., 2006; Frith and Frith, 2006). In both conditions, for answering the question about the expected or unexpected end of the story, participants have to take the perspective of the searching child. It is possible that the activation of the inferior parietal areas during both belief conditions revealed necessary perspective taking.

In summary, we found activation of the dorsal part of the ACC, the right lateral rostral PFC and the right TPJ associated with false belief reasoning. The activation of the dorsal ACC may be related to both the monitoring of action and the reality-independent representation of the protagonist's mental state. The important role of stimulus-independent mental processes during false belief reasoning is additionally supported by the activation of the lateral rostral PFC that seems to be involved in switching attention between stimulus-oriented and stimulus-independent mental processes. The activation of the right TPJ may be associated with the computation of mental representations that create perspective differences, like a person's mental representation of a false belief in contrast to one's own view of reality, and may therefore play a central role in the decoupling mechanism. The conjunction analysis revealed common patterns of activation for true and false belief reasoning, but neither the medial rostral PFC nor the right TPJ, areas, which have been most prominently discussed as neural correlates of mentalizing, appear to be activated in belief reasoning per se. This result may be specific to our belief task as it focused on cognitive components of belief reasoning and excluded other processes like emotion or intention attribution.

References

- Aichhorn, M., Perner, J., Kronbichler, M., Staffen, W., Ladurner, G., 2006. Do visual perspective tasks need theory of mind? *NeuroImage* 30, 1059–1068.
- Amodio, D.M., Frith, C.D., 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev., Neurosci* 7, 268–277.
- Baron-Cohen, S., Leslie, A.M., Frith, U., 1985. Does the autistic child have a “theory of mind”? *Cognition* 21, 37–46.
- Bush, G., Luu, P., Posner, M.I., 2000. Cognitive and emotional influence in anterior cingulate cortex. *Trends Cogn. Sci.* 4, 215–222.
- Castelli, F., Happé, F., Frith, U., Frith, C.D., 2000. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage* 12, 314–325.
- Cavanna, A.E., Trimble, M.R., 2006. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583.
- Den Ouden, H.E.M., Frith, U., Frith, C., Blakemore, S.J., 2005. Thinking about intentions. *NeuroImage* 28, 787–796.
- Fletcher, P.C., Happé, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S.J., Frith, C.D., 1995. Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition* 57, 109–128.
- Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D., Turner, R., 1998. Event-related fMRI: characterizing differential responses. *NeuroImage* 7, 30–40.
- Frith, U., Frith, C.D., 2003. Development and neurophysiology of mentalizing. *Philos. Trans. R. Soc. London, B* 358, 459–473.
- 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 Cogn. Sci.* 7, 77–83.
- Gallagher, H.L., Happé, 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.
- Gilbert, S.J., Frith, C.D., Burgess, H.C., 2005. Involvement of rostral prefrontal cortex in selection between stimulus-oriented and stimulus-independent thought. *Eur. J. Neurosci.* 21, 1423–1431.
- Gilbert, S.J., Simons, J.S., Frith, C.D., Burgess, P.W., 2006a. Performance-related activity in medial rostral prefrontal cortex (Area 10) during low-demand tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 45–58.
- Gilbert, S.J., Spengler, S., Simons, J.S., Steele, J.D., Lawrie, S.M., Frith, C.D., Burgess, P.W., 2006b. Functional specialization within rostral prefrontal cortex (Area 10): a meta-analysis. *J. Cogn. Neurosci.* 18, 932–948.
- Goel, V., Dolan, R.J., 2001. The functional anatomy of humor: segregating cognitive and affective components. *Nat. Neurosci.* 4, 237–238.
- Goel, V., Grafman, J., Sadato, N., Hallett, M., 1995. Modeling other minds. *NeuroReport* 6, 1741–1746.
- Grèzes, J., Frith, C.D., Passingham, R.E., 2004. Inferring false beliefs from the actions of oneself and others: an fMRI study. *NeuroImage* 21, 744–750.
- Henson, R.N.A., Buechel, C., Josephs, O., Friston, K., 1999. The slicetiming problem in event-related fMRI. *NeuroImage* 9, 125.
- Leichnetz, G.R., 2001. Connections of the medial posterior parietal cortex (area 7m) in the monkey. *Anat. Rec.* 263, 215–236.
- 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.* 98, 11832–11835.
- Mitchell, J.P., Banaji, M.R., Macrae, C.N., 2005. General and specific contributions of the medial prefrontal cortex to knowledge about mental states. *NeuroImage* 28, 757–762.
- Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.B., 2005. Valid conjunction inference with the minimum statistic. *NeuroImage* 25, 653–660.
- Perner, J., Aichhorn, M., Kronbichler, M., Staffen, W., Ladurner, G., 2006. Thinking of mental and other representations: the roles of left and right temporo-parietal junction. *Soc. Neurosci.* 1, 245–258.
- Saxe, R., 2006. Why and how to study Theory of Mind with fMRI. *Brain Res.* 1079, 57–65.
- 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., Carey, S., Kanwisher, N., 2004. Understanding other minds: linking developmental psychology and functional neuroanatomy. *Annu. Rev. Psychol.* 55, 87–124.
- Sodian, B., 2005. Theory of mind. The case for conceptual development. In: Schneider, W., Schumann-Hengsteler, R., Sodian, B. (Eds.), *Young Children's Cognitive Development. Interrelationships Among Executive Functioning, Working Memory, Verbal Ability and Theory of Mind*. Erlbaum, Hillsdale, NJ, pp. 95–130.
- Steele, J.D., Lawrie, S.M., 2004. Segregation of cognitive and emotional function in the prefrontal cortex: a stereotactic meta-analysis. *NeuroImage* 21, 868–875.
- Vogeley, K., Fink, G.R., 2003. Neural correlates of the first-person-perspective. *Trends Cogn. Sci.* 7, 38–42.
- Vogeley, K., Bussfeld, P., Newen, A., Hermann, S., Happé, 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., Ciaramidaro, A., Enrici, I., Pia, L., Bara, B.G., 2004. Understanding intentions in social interaction: the role of the anterior paracingulate cortex. *J. Cogn. Neurosci.* 16, 1854–1863.
- Wellman, H.M., Cross, D., Watson, J., 2001. Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684.