Shared and Nonshared Neural Networks of Cognitive and Affective Theory-of-Mind: A Neuroimaging Study Using Cartoon Picture Stories

Lara Schlaffke,¹ Silke Lissek,¹ Melanie Lenz,¹ Georg Juckel,²
Thomas Schultz,^{3,4} Martin Tegenthoff,¹ Tobias Schmidt-Wilcke,¹ and
Martin Brüne^{2*}

¹Department of Neurology, BG-Kliniken Bergmannsheil, Ruhr-University Bochum, Bochum, Germany

²Department of Psychiatry, LWL-University Hospital Bochum, Ruhr-University Bochum, Bochum, Germany

³Department of Computer Science, University of Bonn, Bonn, Germany

⁴Max Planck Institute for Intelligent Systems, Tübingen, Germany

Abstract: Theory of mind (ToM) refers to the ability to represent one's own and others' cognitive and affective mental states. Recent imaging studies have aimed to disentangle the neural networks involved in cognitive as opposed to affective ToM, based on clinical observations that the two can functionally dissociate. Due to large differences in stimulus material and task complexity findings are, however, inconclusive. Here, we investigated the neural correlates of cognitive and affective ToM in psychologically healthy male participants ($n = 39$) using functional brain imaging, whereby the same set of stimuli was presented for all conditions (affective, cognitive and control), but associated with different questions prompting either a cognitive or affective ToM inference. Direct contrasts of cognitive versus affective ToM showed that cognitive ToM recruited the precuneus and cuneus, as well as regions in the temporal lobes bilaterally. Affective ToM, in contrast, involved a neural network comprising prefrontal cortical structures, as well as smaller regions in the posterior cingulate cortex and the basal ganglia. Notably, these results were complemented by a multivariate pattern analysis (leave one study subject out), yielding a classifier with an accuracy rate of more than 85% in distinguishing between the two ToM-conditions. The regions contributing most to successful classification corresponded to those found in the univariate analyses. The study contributes to the differentiation of neural patterns involved in the representation of cognitive and affective mental states of others. *Hum Brain Mapp* 36:29–39, 2015. © 2014 Wiley Periodicals, Inc.

Key words: theory of mind; functional brain imaging; cognitive and affective ToM; multivariate analysis

Conflict of interest: The authors declare no competing financial interests.

Contract grant sponsor: SFB 874 Integration and representation of sensory processes (M.T. and M.L.).

*Correspondence to: Prof. Dr. Martin Brüne, Division of Cognitive Neuropsychiatry and Psychiatric Preventive Medicine, LWL University Hospital, Ruhr-University Bochum,

Alexandrinenstr. 1, D-44791 Bochum, Germany. E-mail: martin.brue@rub.de

Received for publication 20 February 2014; Revised 1 August 2014; Accepted 5 August 2014.

DOI: 10.1002/hbm.22610

Published online 14 August 2014 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

The term “theory of mind” (ToM) concerns the ability to represent own and others’ mental states, such as beliefs, intentions, expectations, knowledge, and emotions. ToM involves the capacity to distinguish self from others and reality from appearance [Leslie, 1987]. For example, an individual can have a representation of the accuracy of another’s knowledge state, that is, that someone else may hold a true or false belief about facts (e.g., the location of an object) or the mental state of a third party (e.g., an intentional deception). The term was originally introduced by primatologists Premack and Woodruff [1978] who examined this cognitive faculty in chimpanzees, and later adopted by developmental psychologists to study ToM in young children. It is now widely accepted that ToM is universal and that the pattern of ToM acquisition during childhood is remarkably similar across cultures, with some differences in developmental speed relating to environmental factors such as parenting and presence or absence of older siblings [reviewed in Brüne and Brüne-Cohrs, 2006]. ToM comprises different levels of complexity, where the true or false belief of an object location is referred to as “second-order” representation, and the inference of the belief that, for instance, a third person has a certain intention or feeling is termed “third-order” representation (“I believe that X assumes that Y intends or feels this or that”; reviewed in [Brüne and Brüne-Cohrs, 2006; Saxe and Powell, 2006].

Research using neuroimaging technique has shown that ToM involves an extended neural network including brain areas in the frontal, temporal, and parietal lobes bilaterally [Amodio and Frith, 2006; Siegal and Varley, 2002]. More precisely, functional brain imaging studies have revealed the activation of cortical midline structures, including the medial prefrontal cortex (mPFC), the anterior cingulate cortex (ACC), and the precuneus and posterior cingulate cortex (PCC), as well as lateral areas of the middle temporal lobes, the temporoparietal junction (TPJ), the superior temporal sulcus, and the temporal poles during the execution of tasks tapping into ToM [Saxe and Kanwisher, 2003; Saxe et al., 2004], whereby a “core network” seems to involve at least the mPFC and the TPJ that is activated across a large range of ToM tasks [Schurz et al., 2014]. With regard to midline structures, previous studies have concluded that the mPFC and the ACC are involved in distinguish self from other and appearance from reality (e.g., deception), and, more generally expressed, in error monitoring and the attribution of saliency [Heatherton et al., 2006; Lissek et al., 2008; Siegal and Varley, 2002].

Moreover, the precuneus seems to be important for the experience of agency and the processing of self-relevant material [Cavanna and Trimble, 2006; Schilbach et al., 2006], but also the representation of others’ thoughts [Saxe and Powell, 2006]. The temporoparietal regions are involved in the representation of others’ thoughts and are consistently activated in a large spectrum of verbal and nonverbal ToM tasks [Apperly et al., 2004; Samson et al., 2004].

More recent studies have aimed at disentangling the activation patterns for cognitive as opposed to affective ToM. Cognitive ToM refers to the ability to represent one’s own and others’ thoughts, intentions, and desires, whereas affective ToM concerns the ability to represent own and others’ emotional states and feelings. This conceptualisation is largely interchangeable with the one distinguishing cognitive and affective empathy [Harari et al., 2010].

In fact, in light of growing evidence for a dissociation of affective versus cognitive ToM coming from studies of child development [Vetter et al., 2014], brain lesions [Shamay-Tsoory et al., 2007], neurodegenerative diseases [Poletti et al., 2012], and psychopathology such as conduct disorder [Sebastian et al., 2012a], schizophrenia, [Shamay-Tsoory et al., 2007] and borderline personality disorder [Harari et al., 2010], brain imaging studies have sought to define the neural correlates of affective and cognitive ToM in a more fine-grained fashion. For example, Völlm et al. [2006], using cartoon picture stories, reported that empathetic perspective-taking (akin to affective ToM) was associated with greater activation of the paracingulate, the anterior and the PCC and the amygdala, whereas cognitive aspects of ToM activated more the lateral orbitofrontal cortex, middle frontal gyrus, cuneus, and superior temporal gyrus compared to the affective condition. Along similar lines, Ochsner et al., [2004] found midline activation during the attribution of emotions to self and others, whereas Hynes et al. [2006] showed orbitofrontal activation for emotional, as compared to cognitive perspective-taking.

When contrasting affective and cognitive ToM directly with one another, a recent study revealed activation in the TPJ and the right cingulate cortex as well as in the left supplementary motor area, whereas the opposite contrast (cognitive vs. affective) did not produce significant differences [Bodden et al., 2013]. Notably, in this study, Bodden et al. [2013] used a relatively simple pictorial ToM task [“Yoni task”; Shamay-Tsoory et al., 2007], where the gaze direction of a cartoon smiley indicates the correct answer. In partial contrast, Corradi-Dell’acqua et al. [2014], utilizing short written vignettes, found overlapping activation during the attribution of emotions and beliefs in the TPJ, but distinguishable (i.e., nonoverlapping) activation in prefrontal regions. Moreover, [Sebastian et al., 2012b] reported, using cartoon vignettes, overlapping neural activation of cognitive and affective facets in the “classic” ToM network in both adolescents and adults, whereby the mPFC was activated only in the affective ToM condition. Furthermore, it was shown that adolescents activated the

Abbreviations

PAP	paper-and-pencil task
PCC	posterior cingulate cortex
SVM	support vector machine
TOM	theory of mind

ventromedial PFC more than adults during affective ToM [Sebastian et al., 2012b]. Along similar lines, Schnell et al. [2011] showed that, when affective ToM was contrasted with visuospatial third-person perspective-taking, affective ToM produced greater activation in the mPFC than visual perspective-taking. These findings have, in part, been corroborated by a recent meta-analysis suggesting that the precuneus and the posterior part of the TPJ are more involved in cognitive ToM (e.g., false belief reasoning), whereas affective ToM elicits more specific activation in inferior frontal areas [Schurz et al., 2014].

Taken together, the aforementioned studies indicate differences in activation patterns between affective and cognitive ToM performance. However, inconsistencies between studies emerged due to profound differences in complexity of the stimulus material. That is, ToM levels varied widely between studies, ranging from visual perspective-taking to more demanding third order representation, the tasks differed also with regard to the sources of information (e.g., emotional contagion vs. cognitive inference), as well as with regard to the material used for the assessment of cognitive and affective ToM, all of which can significantly impact the recruitment of the ToM neural network [Fonagy and Luyten, 2009].

Accordingly, we aimed to study differences in activation between cognitive and affective ToM using an established cartoon-based ToM paradigm [Brüne et al., 2008, 2011a, 2011b; Lissek et al., 2008; Saft et al., 2013]. In particular, we sought to present identical stimulus material for the cognitive, the affective, and the control (“physical”) condition. The paradigm was expanded by doubling the number of scenarios and introducing specific questions about the affective mental states of the cartoon characters, in addition to the established cognitive questions.

We hypothesized that both affective and cognitive ToM performance would be associated with the activation of the ToM “core” neural network. Specifically, we predicted that, in line with previous work and a recent meta-analysis [Schurz et al., 2014], affective ToM would be associated with greater activation of prefrontal structures, particularly the inferior frontal gyrus (IFG), when contrasted directly with cognitive ToM, whereas cognitive ToM would specifically recruit the precuneus and TPJ in comparison to affective ToM.

METHODS

Subjects

Thirty-nine men (mean age 25.9 years; $SD \pm 5.8$) without any psychiatric or neurological diseases were recruited to participate in the present study. We decided to investigate only male participants to keep the group as homogeneous as possible to focus on the differences between affective and cognitive ToM, and at this stage to avoid possible gender effects. The study was conducted in accordance

with the Declaration of Helsinki and approved by the Ethics Committee of the Medical Faculty of the Ruhr-University Bochum, Germany. All participants provided written informed consent.

MRI Protocol

Magnetic resonance imaging (MRI) was performed on a 3.0 Tesla scanner (Philips Achiva 3.2, Best, Netherlands) using a 32-channel head coil. T-1 weighted data sets (TR 8.3 ms, TE 3.8 ms, FOV 256×256 , yielding 220 transversal slices with a voxel size of $1.00 \times 1.00 \times 1.00$ mm and reconstructed to $0.94 \times 0.94 \times 1.00$ mm) were acquired first from all subjects. For the functional analysis 469 dynamic T2* weighted EPI-sequence scans were acquired (TR 3000 ms, TE 35 ms, FOV $224 \times 240 \times 117$ mm, yielding 39 transversal slices with a voxel size of $2.00 \times 2.00 \times 3.00$ mm).

Stimuli

Our original design comprising six ToM cartoon stories [Brüne et al., 2008, 2011b; Lissek et al., 2008; Saft et al., 2013] was expanded by six new cartoons of similar make-up. Thus, four stories depicted a scenario where two characters co-operated with one another, four stories showed a character deceiving another, and four stories depicted two cartoon characters cooperating to deceive a third character. A total of 12 cartoon stories, four pictures each, were presented in a pseudorandomized order in the scanner using fMRI-ready LCD-goggles (VisuaStim, Digital, Resonance Technology Inc., Northridge, CA). For the presentation of the paradigm we used NBS Presentation software (version 9.7). Each story was presented in three conditions with respect to the kind of questions (affective, cognitive, control; 2 questions for each condition). First, the complete story, consisting of four pictures was presented without any questions (for 12 s = 4 scans), and followed by two 12-s episodes with two questions being presented (each for 12 s). Study participants had been instructed to answer the questions mentally, that is, no button press was required (for study design overview see Fig. 1).

Cognitive questions included false-belief questions and questions related to the cartoon character’s intentions, as well as second and third order belief questions (e.g., for a second order type of question: “What does the person in the red shirt think, the one in the blue shirt intends to do?” (see Fig. 2, for an example). The affective questions concerned emotions and feelings of the cartoon characters (e.g., “What does the person in the red shirt feel?”; Fig. 1) whose facial expressions did not reveal their affective states. In the control condition the pictures of the stories were presented in a pseudorandomized order. Questions pertained to properties of objects displayed (e.g., “What color are the trees?”; Fig. 2). The experimental scanning session lasted for approximately 25 min. Participants were

Figure 1.

Example of a cartoon story (cooperation) presented to study subjects. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

familiarized with the study task and study instructions prior to scanning.

Behavioral Measures

To verify whether participants were able to answer the questions correctly, they were given an additional paper-and-pencil task (PAP) after the scanning procedure, which included sequencing the four pictures of each story in chronological order and verbally answering questions sim-

ilar to those presented in the scanner. The PAP-task was evaluated using a score based system. For each cartoon story sequenced correctly, participants received six points, each question answered correctly yielded a maximum score of two points (max. total score 180 points).

In addition, participants completed the Interpersonal Reactivity Index (IRI Davis, [1980], German version from the University of Saarland; [Paulus, 2007]), as a measure of self-reported empathy. For 21 statements in different categories of empathy (perspective taking, fantasy, empathic

Figure 2.

Schematic representation of the study design. Each of the 12 stories was presented 3 times for each condition, leading to a total of 36 pseudo randomized trials. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

TABLE I. Mean scores (given in percent) and standard deviations obtained in the paper-and-pencil version of the Theory of Mind task

Paper-and-pencil theory of mind task			
	Total	Cognitive	Affective
Mean	89.89	96.62	77.07
SD	±3.58	±2.89	±7.80

concerns, and personal distress) participants had to indicate how well these statements describe themselves on a five-level Likert scale ranging from “very well” to “not well.” Averaged test norms for each of the four subdivisions range between 90 and 110 points.

Preprocessing and Statistical Analyses

Functional images were transferred from the scanner to a workstation (Win7, Intel Core i5, 3.2 Ghz) and converted from dicom to nifti (hdr-img pairs) format using MRIconvert 2.0 (Lewis Center for Neuroimaging, University of Oregon). Preprocessing of functional images was performed using SPM8 (Wellcome Department of Cognitive Neurology, University College, London, UK) running under Matlab R2008a, (Mathworks).

The preprocessing steps included realignment for movement correction, spatial normalization to the same stereotactic space (using SPM EPI-template) and spatial smoothing (FWHM: 6 mm).

Within the first level analyses three basic regressors, one for each condition (cognitive, affective, and control), were modeled using a rectangular temporal profile (boxcar with a duration of a single block = 9 TRs) convolved with the hemodynamic response function. Periods when the cartoon stories were presented without questions served as baseline and were not modeled by a specific regressor. Variance estimations were performed within the general linear model [Friston et al., 1995], yielding three β -images, one per condition. The six movement parameters (rotation and translation) were added as nuisance variables to the model. First level analyses also yielded contrast images applying contrast weights (1 -1) to the β -images. We were specifically interested in four contrasts: (i) cognitive ToM > non-ToM; (ii) affective ToM > non-ToM; (iii) cognitive ToM > affective ToM; and (iv) affective ToM > cognitive ToM. The corresponding contrast images were then fed into the second level analyses. Second level analyses (random effects analyses) consisted of one sample t -tests, where voxel wise the null hypothesis was tested that the difference in activation elicited by two conditions was significantly different from zero.

In a first explorative approach (Analysis 1) we applied a very conservative threshold: statistical maps were thresholded at $P < 0.05$ (FWE corrected on a voxel level, see Table III). Due to the very similar nature of the cognitive

tasks involved in the two ToM conditions, both with questions about the social interactions of the cartoon characters, we allowed in a second step (Analysis 2) for the direct comparisons of the two ToM conditions, that is, (iii) cognitive ToM > affective ToM and (iv) affective ToM > cognitive ToM), a slightly more liberal threshold ($P < 0.001$ uncorrected at voxel level followed by $P < 0.05$, FWE corrected at cluster level) to reduce the likelihood false negative findings. Anatomical labeling of brain regions showing significant differences between groups was performed using the SPM8 extension XjView (<http://www.alivelearn.net/xjview8/>). Visualization of results was performed using MRICron (<http://www.cabiatl.com/mricro/mricron/index.html>).

Furthermore, we performed a multivariate analysis (Analysis 3). Specifically, we were interested whether we could train a classifier, a linear support vector machine (SVM), that could differentiate between the two ToM conditions, to identify patterns of brain activation specific to a certain condition. Successful classification provides an additional confirmation of the fact that significant differences exist between the two conditions. For classification, we used a feature vector consisting of 1,24,014 voxels from the differential contrast images, derived from comparing the respective ToM condition to the control state and normalized to zero mean and unit variance. To quantify the degree to which the values of each feature differ between the ToM conditions, F scores were computed for each voxel, defined as

$$F = \frac{(\mu_a - \mu)^2 + (\mu_c - \mu)^2}{\sigma_a^2 + \sigma_c^2}$$

whereby μ is the mean feature value, μ_c , μ_a denote the means for the cognitive and affective conditions, respectively, and σ_c^2, σ_a^2 are the respective variances. This definition assigns a high score to features whose average values differ strongly between the ToM conditions relative to their variance. The impact of voxels less relevant to the classification task was decreased by weighting each feature by its F -score. Classification accuracy was estimated by training a linear SVM with default regularization parameter $C = 1$ on the resulting vectors in a leave-one-subject-out manner; care was taken to exclude test data from feature weighting via the F -score, and to exclude all other scans from the same subject from the training data.

TABLE II. Mean IRI-Scores and standard deviation (given norm-values between 90 and 110)

Interpersonal reactivity index				
	Empathic concern	Perspective taking	Empathic fantasy	Personal distress
Mean	96.43	102.27	91	94.30
SD	±7.47	±6.45	±9.26	±6.89

RESULTS

Paper-and-Pencil-Task

The PAP is a modified written version of the ToM-task, comprising a sequencing task and both affective and cognitive questions. The participants reached a global mean score of 161.79 ± 6.44 of a maximum of 180 (89.89%). For affective questions a mean score of 47.78 ± 4.84 (out of 62; 77.07%) was reached, for cognitive questions participants performed at ceiling level of 114.01 ± 3.4 (out of 118; 96.62%) points. For details see Table I.

Interpersonal Reactivity Index

Participants reached 96.43 ± 7.5 scores for empathic concern, 102.27 ± 6.5 for perspective taking, 91 ± 9.3 for fantasy and 94.3 ± 6.9 scores for personal distress. None of the mean scores fell outside the norm range (for further details see Table II).

Activation Differences Between Affective and Cognitive ToM

Analysis 1: comparisons of the two conditions, affective and cognitive ToM, with the non-ToM control task (contrasts cognitive ToM > non-ToM, affective ToM > non-ToM); (corrected for multiple comparisons on voxel level) revealed substantial similarities. Overlapping activation was found in regions of the network typically activated during theory of mind tasks, namely the medial frontal gyrus, the TPJ, the middle temporal gyrus (BA 21), the supplementary motor area (BA6) and in the precuneus (BA 7). We also observed differences between the two ToM-conditions in the IFG and the expanded precuneus/cuneus region, with cognitive ToM activated larger clusters located in the precuneus and cuneus while affective ToM activating preferentially the IFG (BA45/47). For details see Table III and Figure 3.

Analysis 2: a direct within subject comparison of the two ToM-conditions (FWE corr. $P < 0.05$, cluster level), showing brain regions recruited in one of the ToM condition to a higher extent as compared to the other (contrasts affective ToM > cognitive ToM and cognitive ToM > affective ToM) yielded the following results: cognitive ToM activated a more posterior network including the precuneus, cuneus as well as the middle and superior temporal gyrus when contrasted with affective ToM. Conversely, affective ToM recruited an anterior network consisting of the superior and middle frontal gyrus and the IFG bilaterally as well as the caudate bilaterally, as compared to cognitive ToM (see Table III and Fig. 4).

fMRI Results—Multivariate Analyses

Analysis 3: in the leave-one-subject-out classification experiment, 66 out of 76 conditions (86.8%, chance level at

50%) were classified correctly as being affective or cognitive. A binomial test rejected ($P < 0.01$) the null hypothesis that this result was obtained by chance. Visual inspection of the weight vector $\mathbf{w}$, used by the SVM to distinguish between the conditions, indicates that multivariate analysis relied on regions similar to the ones found by univariate testing. In particular, greater activation in the inferior and superior frontal gyrus contributed to classification as affective ToM, while activation in the cuneus and middle temporal gyrus contributed to classification as cognitive ToM (Fig. 5).

DISCUSSION

The present study sought to examine the neural correlates of affective versus cognitive ToM, based on previous studies that have produced somewhat inconclusive results, which, in part, might be related to differences in stimulus material for the two ToM conditions. Here, we showed partially overlapping and partially distinguishable brain activations when thinking about others' cognitive, as contrasted to affective mental states. Specifically we found, in line with previous work [Schnell et al., 2011; Schurz et al., 2014; Sebastian et al., 2012a; Ochsner et al., 2004; Völlm et al., 2006], that cognitive ToM was associated with greater activation in the precuneus, cuneus as well as the middle and superior temporal gyrus, when contrasted with affective ToM. Conversely, in the affective ToM condition participants showed greater activation in the superior and middle frontal gyrus as well as the IFG [Schurz et al., 2014]. A classifier yielded a high accuracy rate distinguishing between the two conditions, providing additional evidence that activations differ significantly between the two ToM conditions and suggesting that brain activation patterns could potentially be used to predict a participant's analysis approach to a given stimulus. This finding is particularly important in view of the fact that identical stimulus material was used for the two ToM-conditions, which differed only with regard to the questions presented on the screen.

Neuropsychological as well as imaging studies have revealed that a useful distinction can be made between the representation of cognitive mental states (beliefs, desires, intentions, etc.) and affective mental states (feelings and emotions), with overlapping, but also distinguishable neural representations [Bodden et al., 2013; Corradi-Dell'acqua et al., 2014; Hynes et al., 2006; Ochsner et al., 2004; Shamay-Tsoory et al., 2007]. Arguably, affective ToM is closely related to empathetic perspective-taking, suggesting that the representation of such mental states involves prefrontal cortical midline structures, as shown in previous studies [e.g., Sebastian et al., 2012b]. In contrast, cognitive ToM, more than affective ToM, may necessitate the distinction between appearance from reality (represented in the PCC/precuneus), at least if deception is involved, as was the case with our stimulus material [Lissek et al., 2008].

TABLE III. Contrasts of the ToM Task ($n = 39$) analysis 1 (FWE corrected on voxel level) and analysis 2 (FWE corrected on cluster level). Clusters with a minimum size of 20 voxels are reported. Clusters with asterisks indicate significance after FWE correction on voxel level (analysis 1)

		Analysis 1						Analysis 2																	
		FWE corrected on voxel level p<0.05						FWE corrected on cluster level p<0.05 (p<0.001 uncorrected on voxel level)																	
		Affective ToM> Control						Cognitive ToM> Control						Affective ToM> cognitive ToM						Cognitive ToM> affective ToM					
		X	y	z	t-value	Voxel #	BA	X	y	z	t-value	Voxel #	BA	X	y	z	t-value	Voxel #	BA	X	y	z	t-value	Voxel #	BA
Frontal lobe	Superior Frontal Gyrus	-4	12	55	11.21	581	6	-4	14	58	6.45	22	6	-2	30	40	5.86	in 4157	6						
		-4	16	49	9.62	in 581	8																		
	Medial Frontal Gyrus	-8	56	34	8.54	in 581	8																		
		-44	18	49	8.82	248	6	-46	12	52	9.01	190	6												
	Inferior Frontal Gyms	-52	22	10	10.48	554	45								-52	18	4	6.48	in 4157	45					
Temporal Lobe	Superior Temporal Gyrus	-58	-60	19	11.14	1172	22																		
								-52	-64	19	8.51	1711	39												
	Middle Temporal Gyrus	-52	-36	-5	9.39	in 1172	21	52	-48	16	11.51	1220	39												
								-60	-38	-2	9.21	in 1711	21												
Occipital Lobe	Inferior Temporal Gyrus	60	-8	-20	6.65	26	21	50	-64	13	9.83	in 1220	21												
								52	-14	20	7.41	68	21												
	Superior Occipital Gyrus							-56	-6	-23	8.96	101	21												
	Middle Occipital Gyrus	52	-76	1	12.32	553	39																		
		30	-98	4	8.02	27	18																		
		14	-106	10	8.26	34	18																		
	Inferior Occipital Gyrus	42	-76	-8	8.28	in 553	19																		
	Cuneus							-48	-80	1	10.51	in 1711	18												
Parietal Lobe	Precuneus	-4	-54	37	9.27	337	7	-2	-58	40	13.17	1173	7												
Limbic Lobe																									
Sub Lobar	Posterior cingulate																								
	Caudate

Figure 3.

Shared and nonshared neural networks of cognitive and affective ToM (Analysis 1). Statistical parameter maps of shared and nonshared neural networks of affective and cognitive theory of mind. Contrast between affective and control condition (affective > control) is

shown in blue, contrast between cognitive and control condition (cognitive > control) is shown in green, overlap in turquoise. TPJ = temporoparietal junction. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

In support of this assumption, a particular strong activation was found in the precuneus/cuneus region when participants thought about others' cognitive mental states, but not affective mental states. This is interesting, because the precuneus seems to be important for the experience of agency and the processing of self-relevant material [Cavanna and Trimble, 2006; Schilbach et al., 2006], and particularly the representation of others' thoughts [Saxe and Powell, 2006]. In one of our previous studies [Lissek et al., 2008], we found precuneus activation in all contrasts involving cooperation, deception, and cooperation/deception compared with the non-ToM condition, similar to what Sommer et al. [2007] reported for false and true belief reasoning. Along these lines, in a positron emission tomography study Ruby and Decety [2001] found bilateral precuneus activation when participants were taking a third person perspective compared to a first person perspective.

Notably, affective ToM specifically involved the recruitment of smaller clusters located in the PCC and striatum. PCC activation has been repeatedly been demonstrated for ToM and/or empathy tasks [Fletcher et al., 1995; Goel et al., 1995; Saxe and Powell, 2006]. However, our finding of stronger mPFC activation in the affective ToM condition (as contrasted with the cognitive ToM condition) is most consistent with Völlm et al. [2006], Ochsner et al. [2004] as well as Sebastian et al. [2012b], who also reported stronger mPFC activation during the processing of affective mental states compared to cognitive mental states.

Another site that showed a strong activation associated with affective ToM was the left IFG, suggesting that this activation site might be even more specific to affective ToM. The IFG overlaps with the inferolateral frontal cortex

region identified by Abu-Akel as crucial for the automatic matching of others' and own mental states [Abu-Akel, 2003]. Although speculative, affective ToM, as designed in our task, could have elicited emotional contagion, possibly processed via the action of mirror neurons located in the IFG [Iacoboni and Dapretto, 2006; Kilner et al., 2006]. In any event, our findings are largely consistent with a recent meta-analysis suggesting differential activation for cognitive versus affective components of ToM [Schurz et al., 2014].

Another issue that warrants a more detailed discussion pertains to the activation of the basal ganglia in the affective, as contrasted to the cognitive ToM condition. This finding is strikingly similar to what Bodden et al. [2013] reported, who reported activation clusters located in the caudate nucleus when contrasting affective ToM versus a physical control condition, whereas this region did not specifically distinguish between affective and cognitive ToM [Bodden et al., 2013]. However, this similarity is noteworthy in light of evidence for ToM deficits in diseases associated with basal ganglia dysfunction such as Parkinson's disease [Bodden et al., 2010] and Huntington's disease [Brüne et al., 2011b], whereby preclinical Huntington gene carriers do not show aberrant activation patterns in a ToM task compared to controls [Saft et al., 2013]. Interestingly, and in line with our hypothesis, Bodden et al. [2013] suggested that the basal ganglia play a role in the mirror neuron system [Alegre et al., 2010], supporting the notion that the mirror neuron system could be involved in affective ToM. However, this remains to be fully explored using study designs that more explicitly tap into the matching of cognitive and affective mental states.

Figure 4.

Differences in brain activation between cognitive and affective ToM (Analysis 2). Statistical parameter maps of the within subject direct comparisons of the two ToM conditions. Contrast affective > cognitive ToM is shown in red in the top-panel, contrast cognitive > affective ToM is shown in yellow in the bottom panel. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

Figure 5.

Weighting factors of the SVM (Analysis 3). Coefficients of the weighting vectors of the SVM. In regions shown in blue, a higher activation led to cognitive classification, in regions shown in red, a higher activation led to an affective classification. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

The present study has a couple of limitations which need to be mentioned. First, cognitive and affective ToM were examined in a group of psychologically healthy young male participants, which precludes generalisations across the lifespan and gender. Second, we did no control for differences in important personality variables (such as the “Big Five”). However, we assume that personality traits were equally distributed across the group inducing no systematic bias to the study population. Moreover, we did not see any correlation with the PAP-task, which was probably due to the fact that participants performed at ceiling level in the task. Finally, we cannot rule out differences in activation due to differences in task complexity. That is, while cognitive ToM questions involved first to third order representations, the affective questions involved only second-order representation.

In summary, our study lends further support to previous observations indicating that affective and cognitive aspects of ToM are processed in shared, but in part also distinct neural networks. Future studies will need to investigate in more detail connections and interactions between the aforementioned brain regions. Understanding the distinction between cognitive and affective components of ToM including their underlying neural networks will also help improve insights into the nature of neuropsychiatric and psychopathological conditions in which these two aspects have fallen functionally apart.

REFERENCES

- Abu-Akel A (2003) A neurobiological mapping of theory of mind. *Brain Res Rev* 43:29–40.
- Alegre M, Rodríguez-Oroz MC, Valencia M, Pérez-Alcázar M, Guridi J, Iriarte J, Obeso JA, Artieda J (2010) Changes in subthalamic activity during movement observation in Parkinson's disease: is the mirror system mirrored in the basal ganglia? *Clin Neurophysiol* 121:414–425.
- Amodio DM, Frith CD (2006) Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci* 7:268–277.
- Apperly IA, Samson D, Chiavarino C, Humphreys GW (2004). Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. *J Cogn Neurosci* 16: 1773–1784.
- Bodden ME, Kübler D, Knake S, Menzler K, Heverhagen JT, Sommer J, Kalbe E, Krach S, Dodel R (2013) Comparing the neural correlates of affective and cognitive theory of mind using fMRI: Involvement of the basal ganglia in affective theory of mind. *Adv Cogn Psychol* 9:32–43.
- Bodden ME, Mollenhauer B, Trenkwalder C, Cabanel N, Eggert KM, Unger MM, Oertel WH, Kessler J, Dodel R, Kalbe E (2010) Affective and cognitive Theory of Mind in patients with parkinson's disease. *Parkinsonism Relat Disord* 16:466–470.
- Brüne M, Brüne-Cohrs U (2006) Theory of mind-evolution, ontogeny, brain mechanisms and psychopathology. *Neurosci Biobehav Rev* 30:437–455.
- Brüne M, Lissek S, Fuchs N, Witthaus H, Peters S, Nicolas V, Juckel G, Tegenthoff M (2008) An fMRI study of theory of mind in schizophrenic patients with “passivity” symptoms. *Neuropsychologia* 46:1992–2001.
- Brüne M, Blank K, Witthaus H, Saft C (2011a) “Theory of mind” is impaired in Huntington's disease. *Mov Disord* 26:671–678.
- Brüne M, Ozgürdal S, Ansorge N, von Reventlow HG, Peters S, Nicolas V, Tegenthoff M, Juckel G, Lissek S (2011b) An fMRI study of “theory of mind” in at-risk states of psychosis: comparison with manifest schizophrenia and healthy controls. *Neuroimage* 55:329–337.
- Cavanna AE, Trimble MR (2006) The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129:564–583.
- Corradi-Dell'acqua C, Hofstetter C, Vuilleumier P (2014) Cognitive and affective theory of mind share the same local patterns of activity in posterior temporal but not medial prefrontal cortex. *Soc Cogn Affect Neurosci* 9:1175–1184.
- Davis MH (1980) Interpersonal Reactivity Index 8. *JSAS Cat Sel Doc Psychol* 10:85.
- Fletcher PC, Happé F, Frith U, Baker SC, Dolan RJ, Frackowiak RS, Frith CD (1995) Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition* 57:109–128.
- Fonagy P, Luyten P (2009) A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Dev Psychopathol* 21: 1355–1381.
- Friston, KJ, Holmes, AP, Worsley, KJ, Poline, JP, Frith, CD, Frackowiak, RSJ (1995) Statistical parametric maps in functional imaging: A general linear approach. *Hum Brain Map* 2: 189–210.
- Goel V, Grafman J, Sadato N, Hallett M (1995) Modeling other minds. *Neuroreport* 6:1741–1746.
- Harari H, Shamay-Tsoory SG, Ravid M, Levkovitz Y (2010) Double dissociation between cognitive and affective empathy in borderline personality disorder. *Psychiatry Res* 175:277–279.
- Heatherton TF, Wyland CL, Macrae C N, Demos K E, Denny BT, Kelley WM (2006) Medial prefrontal activity differentiates self from close others. *SCAN* 1:18–25.
- Hynes CA, Baird AA, Grafton ST (2006) Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia* 44:374–383.
- Iacoboni M, Dapretto M (2006) The mirror neuron system and the consequences of its dysfunction. *Nat Rev Neurosci* 7:942–951.
- Kilner JM, Marchant JL, Frith CD (2006) Modulation of the mirror system by social relevance. *Soc Cogn Affect Neurosci* 1:143–148.
- Leslie, A. M. (1987). Pretense and Representation: The Origins of Theory of Mind. *Psychol Rev* 94: 412–426.
- Lissek S, Peters S, Fuchs N, Witthaus H, Nicolas V, Tegenthoff M, Juckel G, Brüne M (2008) Cooperation and deception recruit different subsets of the theory-of-mind network. *PLoS One* 3: e2023.
- Ochsner KN, Knierim K, Ludlow DH, Hanelin J, Ramachandran T, Glover G, Mackey SC (2004) Reflecting upon feelings: an fMRI study of neural systems supporting the attribution of emotion to self and other. *J Cogn Neurosci* 16:1746–1772.
- Paulus C (2007) Der Saarbrücker Persönlichkeitsfragebogen zur Empathie (IRI). Available at: <http://www.uni-saarland.de/fak5/ezw/personal/paulus/empathy/SPF.htm> in July 2012.
- Poletti M, Enrici I, Adenzato M (2012) Cognitive and affective Theory of Mind in neurodegenerative diseases: Neuropsychological, neuroanatomical and neurochemical levels. *Neurosci Biobehav Rev* 36:2147–2164.

- Premack D, Woodruff G (1978) Does the chimpanzee have a theory of mind? *Behav Brain Sci* 1:515–526.
- Ruby P, Decety J (2001) Effect of subjective perspective taking during simulation of action. *Nat Neurosci* 4:546–550.
- Saft C, Lissek S, Hoffmann R, Nicolas V, Tegenthoff M, Juckel G, Brüne M (2013) Mentalizing in preclinical Huntington's disease: An fMRI study using cartoon picture stories. *Brain Imaging Behav* 7:154–162.
- Samson D, Apperly IA, Chiavarino C, Humphreys GW (2004) Left temporoparietal junction is necessary for representing someone else's belief. *Nat Neurosci* 7:499–500.
- Saxe R, Powell LJ (2006) It's the thought that counts: Specific brain regions for one component of theory of mind. *Psychol Sci* 17: 692–699.
- Saxe R, Carey S, Kanwisher N (2004) Understanding other minds: linking developmental psychology and functional neuroimaging. *Ann Rev Psychol* 55:87–124.
- 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.
- Schilbach L, Wohlschlaeger AM, Kraemer NC, Newen A, Shah NJ, Fink GR, Vogeley K (2006) Being with virtual others: Neural correlates of social interaction. *Neuropsychologia* 44:718–730.
- Schnell K, Bluschke S, Konrad B, Walter H (2011) Functional relations of empathy and mentalizing: an fMRI study on the neural basis of cognitive empathy. *NeuroImage* 54:1743–1754.
- Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014) Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci Biobehav Rev* 42:9–34.
- Sebastian CL, McCrory EJ, Cecil CA, Lockwood PL, Brito SA, Fontaine NM, Viding E (2012a) Neural responses to affective and cognitive theory of mind in children with conduct problems and varying levels of callous-unemotional traits. *Arch Gen Psychiatry* 69:814–822.
- Sebastian, CL, Fontaine, NM, Bird, G, Blakemore, SJ, Brito, SA, McCrory, EJ, Viding, E (2012b). Neural processing associated with cognitive and affective Theory of Mind in adolescents and adults. *SCAN* 7:53–63.
- Shamay-Tsoory SG, Aharon-Peretz J, Levkovitz Y (2007) The neuroanatomical basis of affective mentalizing in schizophrenia: comparison of patients with schizophrenia and patients with localized prefrontal lesions. *Schizophr Res* 90:274–283.
- Siegal M, Varley R (2002) Neural systems involved in "theory of mind". *Nat Rev Neurosci* 3:463–471.
- Sommer M, Döhl K, Sodian B, Meinhardt J, Thoermer C, Hajak G (2007) Neural correlates of true and false belief reasoning. *Neuroimage* 35:1378–1384.
- Vetter NC, Weigelt S, Döhl K, Smolka MN, Kliegel M (2014) Ongoing neural development of affective theory of mind in adolescence. *Soc Cogn Affect Neurosci* 9:1022–1029.
- Völlm B, Taylor N, Richardson P, Corcoran R, Stirling J, McKie S, Deakin JF, Elliott R (2006) Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *NeuroImage* 29:90–98.