Review

Fractionating theory of mind: A meta-analysis of functional brain imaging studies

Matthias Schurz\textsuperscript{a,b,\*}, Joaquim Radua\textsuperscript{c,d}, Markus Aichhorn\textsuperscript{a,b}, Fabio Richlan\textsuperscript{a,b}, Josef Perner\textsuperscript{a,b}

\textsuperscript{a} Department of Psychology, University of Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria
\textsuperscript{b} Center for Neurocognitive Research, University of Salzburg, Hellbrunnerstr. 34, 5020 Salzburg, Austria
\textsuperscript{c} Institute of Psychiatry, King’s College London, De Crespigny Park, London SE5 8AF, United Kingdom
\textsuperscript{d} FIDMAG Germanes Hospitalaries Research Unit, CIBERSAM, Sant Boi de Llobregat, Barcelona, Spain

\begin{abstract}
We meta-analyzed imaging studies on theory of mind and formed individual task groups based on stimuli and instructions. Overlap in brain activation between all task groups was found in the mPFC and in the bilateral posterior TPJ. This supports the idea of a core network for theory of mind that is activated whenever we are reasoning about mental states, irrespective of the task- and stimulus-formats (Mar, 2011). In addition, we found a number of task-related activation differences surrounding this core-network. ROI based analyses show that areas in the TPJ, the mPFC, the precuneus, the temporal lobes and the inferior frontal gyri have distinct profiles of task-related activation. Functional accounts of these areas are reviewed and discussed with respect to our findings.
\end{abstract}

© 2014 Published by Elsevier Ltd.
3.2. Region of interest (ROI) analyses

3.2.1. Temporo-parietal ROIs

3.2.2. Cortical midline ROIs

3.2.3. Fronto-temporal ROIs

4. Discussion

4.1. The role of conceptual variance for meta-analysing theory of mind

4.2. Contribution of the present meta-analysis

4.3. Core network for theory of mind

4.4. Areas for theory of mind

4.4.1. Temporo-parietal junction

4.4.2. Medial prefrontal cortex

4.4.3. Precuneus

4.4.4. Anterior temporal lobes

4.4.5. Inferior frontal gyrus

4.5. Limitations

5. Conclusion

Appendix A. Supplementary data

References

1. Introduction

A central component of human social cognition is the ability to attribute mental states to ourselves and others. This ability is often referred to as ‘mentalizing’, ‘mindreading’ or ‘theory of mind’. For nearly two decades, functional imaging has been used to study how this ability is implemented in the brain. To date, hundreds of empirical studies on this topic can be found in the literature. In the present study, we meta-analyze these imaging findings by forming different task groups based on the stimuli and instructions used.

Traditionally, the literature has been summarized in narrative reviews (e.g., Saxe, 2006; Frith and Frith, 2006). Some reviews proposed that all sorts of theory of mind tasks consistently engage a particular brain network (e.g., Amodio and Frith, 2006; Frith and Frith, 2006; Mitchell, 2009). This network includes the medial prefrontal cortex (mPFC), parts of precuneus and posterior cingulate cortex, the temporo-parietal junction (TPJ) and the posterior superior temporal sulcus (pSTS) bilaterally. Less frequently, the bilateral anterior temporal lobes and areas in the amygdala have also been associated with this network (Frith and Frith, 2006; Mar, 2011).

The common feature of all tasks activating this network is that they require thinking about the mental states of other persons (Frith and Frith, 2006). For example, Gallagher and Frith (2003, p. 78) noted that functional imaging studies on mentalizing have ‘... demonstrated remarkable consistent results despite using multimodal and diverse cognitive paradigms, such as verbal and non-verbal or on- and off-line tasks. Similarly, Mitchell (2009, p. 1310) noted: ‘... studies have established a small and highly reliable network of regions that is preferentially engaged when perceivers mentalize about the mind of others.’ The author added ‘... researchers observed this pattern of activation regardless of whether mental inferences are prompted by stories, cartoons... In the context of competitive and economic games... or by task instructions to think about a specific person’s mind.’

At the same time, many researchers say that particular brain areas subserve particular subprocesses of theory of mind (e.g., Frith and Frith, 1999; Saxe et al., 2004a, 2004b; Gallagher and Frith, 2003; Amadio and Frith, 2006; Gobbini et al., 2007; Perner and Leekam, 2008). Moreover, it has been proposed that some subprocesses of theory of mind are preferentially engaged in some tasks and not in others (e.g., Aichhorn et al., 2006; Bahnemann et al., 2010; Gobbini et al., 2007; Perner and Leekam, 2008). For example, Gobbini et al. (2007) observed that theory of mind tasks that focus on false beliefs tend to activate the TPJ more dorsally than tasks without false beliefs (e.g., social animations and point-light movement perception), which activates more ventrally in the pSTS.

In the recent years, quantitative meta-analyses have become a tool for validating and extending the findings of narrative imaging reviews. They offer a statistical procedure to objectively summarize imaging findings in the form of brain-maps of convergent activation (e.g., Eickhoff et al., 2009; Radua et al., 2012). Several groups (e.g., Bzdok et al., 2012; Decety and Lamm, 2007; Spreng et al., 2009) performed ‘pooled’ meta-analyses on theory of mind, i.e., putting together all published imaging studies that used a term like ‘theory of mind’ in their keywords. Findings of these meta-analyses show the same large-scale network of areas that has been described in narrative reviews. A methodological strength of pooled meta-analyses is that they have high statistical power to detect findings, as they are based on a large sample of all available studies. Large variability, however, is also a drawback of these approaches, as they put together studies using different stimulus material and control conditions, as well as studies with variable methodological quality and rigor. In a conceptually diverse field such as theory of mind research, this will produce a lot of noise. For example, Mar (2011) noted that ‘Differences in approach to studying ToM may explain some of the current disagreement regarding which areas are key for mentalizing’ (p. 117). Up to now, only one review has looked at task-related differences in activation for theory of mind with a quantitative method (Van Overwalle, 2009). The author performed a region of interest based meta-analysis on social cognition, including the domain of theory of mind. A limitation of van Overwalle’s (2009) approach is that findings of his method are bound to the scope of the a priori defined regions of interest. For example, the TPJ was analyzed with one single region of interest, spanning from pSTS (at z = 3) up to the dorsal TPJ (located in the inferior parietal lobe, at z = 39). The author observed convergent activation in this region for a wide variety of theory of mind tasks. For example, consistent activation was found in the TPJ for false belief tasks as well as for social animations. This observation stands in contrast to what was proposed by Gobbini et al. (2007) and other researchers (e.g., Aichhorn et al., 2006; Bahnemann et al., 2010; Perner and Leekam, 2008).

In the present study, we performed a review of theory of mind imaging studies, and sorted them into task groups that had comparable stimulus-material, instructions and control conditions. This reduces the problem of pooling across studies which differ on conceptual and methodological aspects, and offers a more-fine grained picture of the evidence. In addition, we performed a whole-brain meta-analysis on a voxel-by-voxel basis, which allowed us to look at local differences in more detail than earlier works (e.g., Van Overwalle, 2009). Our meta-analysis aims at providing a systematic and quantitative evaluation of two previous claims. First, we...
wanted to evaluate to what extent different theory of mind tasks consistently activate a common set of brain areas (e.g., Amodio and Frith, 2006; Frith and Frith, 2006; Mitchell, 2009). Second, we wanted to evaluate if some sub-areas of the theory of mind brain network are preferentially engaged by some tasks and not by others (Aichhorn et al., 2006; Bahnemann et al., 2010; Gobbini et al., 2007; Perner and Leekam, 2008).

2. Methods

2.1. Literature search and study selection

The literature reviewed in our meta-analysis was retrieved by several strategies. First, we adopted the literature sample of a recent large-scale meta-analysis (Mar, 2011). The author categorized studies into story-based versus non-story based tasks, and looked at meta-analytic overlaps between them. In contrast, the present study seeks to meta-analyze brain activation for individual groups formed on the basis of a certain type of task, i.e., studies that have presented comparable sorts of stimuli and instructions. Mar’s meta-analysis contained literature on theory of mind published between 1995 and 2010 with the following strategy: Mar’s (2011) performed a key-word search in the databases PubMed, Science Citation Index, and PsychInfo. The criteria were the following: studies were selected if (i) they included one of the key-words ‘neuroimaging’ or ‘fMRI’ or ‘PET’ and (ii) one of the key-words ‘theory-of-mind’ or ‘mentalizing’ or ‘mindreading’. We updated Mar’s (2011) literature sample by replicating the author’s search for the literature published between 2010 and 2013 (up to January 2013). Based on the literature sample retrieved in these two steps, we attempted to identify task groups representing frequently used experimental paradigms to study theory of mind. A task group had to contain studies which used similar stimuli and presentation-modes, similar task-instructions, and results from comparable contrasts. As we are not aware of exact criteria for the minimum sample size in imaging meta-analysis, we ensured that our task groups had samples within the typical range found in the literature. We refer to a recent summary article of 94 published imaging meta-analyses (David et al., 2013). This paper reports that the interquartile-range of the number of participants included in imaging meta-analysis lies between 123 and 355, and the IQR for the number of original studies included lies between 8 and 20. All our task groups had sample-sizes well within that range. After a preliminary list of task groups was defined, we searched through multiple existing literature reviews on theory of mind and related topics to find additional studies falling into these groups (Bzdok et al., 2012; Denny et al., 2012; Murray et al., 2012; Perner and Leekam, 2008; Spreng et al., 2009; Van Overwalle, 2009; Van Overwalle and Baetens, 2009).

We then applied a number of methodological selection-criteria to the body of literature identified by our search (see e.g., Radua et al., 2012). Studies were only selected if they had performed a whole brain analysis and reported activation coordinates in standard space (MN1 or Talairach). Data from clinical samples were removed. We ensured that the same threshold throughout the whole brain was used within each included study, in order to avoid biases toward liberal thresholded brain regions. Note that this does not mean that different studies should employ the same threshold. We also ensured that all activation maps entering our meta-analysis refer to a basic process of reasoning about/processing mental states. Therefore, we only selected studies which contrasted a task clearly referring to mental state processing versus a task which clearly did not. This excluded several study types, for example studies which compared inferences about others’ versus own mental states or which did not engage mental states in any condition. If a study reported more than one contrast, we selected the one best corresponding to contrasts reported in other studies of a task group. In total, we meta-analyzed 757 activation foci reported from 73 studies that tested 1241 participants. Our meta-analysis contained six different task groups: (i) false belief, including 112 activation foci from 15 studies with 259 participants, (ii) trait judgment tasks, 111 foci, 15 studies, 253 participants, (iii) strategic games, 68 foci, 9 studies, 162 participants, (iv) social animations, 203 foci, 14 studies, 224 participants, (v) mind in the eyes tasks, 146 foci, 10 studies, 185 participants and (vi) rational actions, 117 foci, 10 studies, 158 participants.

2.2. Meta-analytic methods

2.2.1. Effect-size signed differential mapping

We used Effect-Size Signed Differential Mapping (ES-SDM) software, version 2.31 for meta-analysis (Radua et al., 2012; www.sdmproject.com). SDM is based and improves upon the positive features from existing peak probability methods for meta-analysis, such as Activation Likelihood Estimation (ALE, Eickhoff et al., 2009) or Multilevel Kernel Density Analysis (MKDA, Wagner et al., 2007). ES-SDM uses standard effect size and variance-based meta-analytic calculations. Based on the reported t-values and the sample size of a study, ES-SDM creates a map of effect-sizes (Hedge’s g values) and a map of variance, with the latter being derived from the distribution of effect-sizes and the sample size of the study. Effect-sizes are calculated for those voxels containing a peak that is reported in the results table of an original study. For the remaining voxels, an effect-size is estimated depending on the distance to close peaks (<20 mm) by means of an unnormalized Gaussian kernel. In the present analysis, we used the recommended Gaussian kernel with a FWHM of 20 mm. A validation study which compared the results of coordinate based ES-SDM meta-analysis to the results of a standard voxel-wise GLM analysis of the same original data (Radua et al., 2012) found that this FWHM provided an optimal balance between sensitivity and specificity. For statistical analysis, all foci were transformed to Talairach space which is the native space of the software, by using the matrix transformations proposed by Lancaster et al. (2007).

We calculated a mean analysis for each task group. Calculation of the meta-analytic mean map is implemented by a random-effects model in which each study is weighted by the inverse of the sum of its variance plus an estimate of between-study heterogeneity. The latter is obtained by the DerSimonian–Laird method (DerSimonian and Laird, 1986). This approach enables studies with larger sample size or lower variability to contribute more, and that effects are assumed to randomly vary between samples. Statistical significance of the mean map was assessed by a permutation test that randomizes the location of the voxels within the standard SDM gray matter template. Thus, the null hypothesis of these permutation tests is that effect-sizes are randomly distributed throughout the brain. 100 random maps were generated with the same number of input foci as included in the to-be-tested map. Finally, the meta-analytic maps were thresholded using a voxel-level (height) threshold of p<0.005 (uncorrected) and a cluster-level (extent) threshold of 10 voxels. This uncorrected threshold was found to optimally balance sensitivity and specificity, and to be an approximate equivalent to a corrected threshold of p<0.05 in original neuroimaging studies (Radua et al., 2012). For convenience, we report all activations in MNI-space.

2.2.2. Complementary analyses

We used the heterogeneity analysis in ES-SDM to check whether the brain areas found by our mean analysis showed strong variability across studies. ES-SDM calculates a Q-statistic based on the between study variance in effect-size estimates for a given area. Heterogeneity can then be tested for significance by determining if
the observed between-study variance for a given area is larger than that resulting from sampling error alone. For convenience, heterogeneity values are converted into standard z values in ES-SDM.

Systematic whole-brain voxel-based jackknife sensitivity analysis was conducted. This consists of repeating the main statistical analysis for each study in the meta-analysis once, and discarding a different study each time. This means, we removed one study at a time, performed an analysis on the remaining studies, replaced the study and repeated this process. The rationale of this test is that if a previously significant brain region remains significant in all or most of the combinations of studies, it can be concluded that this finding is highly replicable.

2.2.3. Overlap analyses

We used three methods to determine brain regions that were activated by all task groups. First, we used conjunction analysis (Nichols et al., 2005), which calculates a simple overlap between regions that were found statistically significant in the individual meta-analyses. This was done with the ‘image calculator’ utility in SPM8 (www.filion.ucl.ac.uk). This procedure shows findings which are individually significant in all six meta-analyses at our default threshold of p < .005 uncorrected, z > 1, and minimum cluster extent 10 voxels.

However, the conjunction approach may be biased toward a false impression of selectivity. This is because the number of studies in each of the task-specific meta-analyses is much lower than the number of studies in the pooled analysis, thus reducing the power of the individual analyses. To address this problem, we carried out a permutation based overlap analysis. Here, we determined the overlap between two halves of the pooled-meta analysis sample, each consisting of the studies from three different task groups. This procedure was repeated ten times to create every permutation for assigning six task groups to two sample-halves. Carrying out conjunction analyses between the sample-halves yielded ten maps of overlap, which were then again merged by conjunction analysis, giving a final map of overlap between task groups.

The third way to determine overlaps between task groups was a multimodal meta-analysis as described by Radua et al. (2013). Whereas the conjunction analysis determines the overlap between binary thresholded maps, multimodal meta-analysis determines overlap based on the actual p-values from the original meta-analytic maps. Thereby, the method takes into account that p-values from the unimodal meta-analyses may be estimated with some degree of error. Significance is determined by an adjusted U statistic based on the convenient null-hypothesis that not a single one of the meta-analyses shows activation in an area (for more details, see Radua et al., 2013). The multimodal meta-analysis was initially created for the case of two meta-analyses (e.g., Radua et al., 2012; Cooper et al., 2013) and it was noted that the method should be conducted with caution in the case of more than two meta-analyses, as it may become too liberal. Conjunction analysis, on the other hand, becomes increasingly conservative in the case of more than two individual meta-analyses.

2.2.4. ROI analyses

We focused on ROI analyses to characterize the task-specific patterns of activation underlying the main findings of our pooled-analysis. Therefore, we placed ROIs at the locations of peak voxels from our pooled analysis (highlighted in bold in Table 3). For large clusters (temporo-parietal clusters, medial prefrontal cluster), we additionally placed ROIs at local subpeaks, according to the following criteria: (i) Centers of all ROIs had a minimum distance of 20 mm, which corresponds to the FWHM of meta-analytic smoothing. (ii) ROIs were placed in different functional-anatomical areas or subdivisions of areas (e.g., the TPJ), as described by brain atlases. For meta-analytic clusters that were found both in the left and right hemisphere, we started with ROI selection on the right side, and then chose closest possible peak or sub-peak on the homologous left coordinate. For the TPJ and the mPFC we relied on recent connectivity-based parcellation atlases (Mars et al., 2011, 2012; Sallet et al., 2013) for labeling our results and for selection of ROIs. These atlases subdivide the TPJ and the mPFC based on patterns of structural and functional connectivity (for details see Section 2.3).

For locating other ROIs in our analysis, such as prefrontus and fronto-temporal ROIs, we relied on classical macroanatomic atlas boundaries (Talairach Daemon). The only exceptions to this procedure were our ROIs in the anterior temporal lobes. We selected the anterior temporal ROIs because they were located close to the temporal poles, for which we could not find proper local subpeaks. For one finding from our pooled meta-analysis, the left fusiform gyrus, we did not analyze ROI data because of its closeness to another ventral temporal ROI (left middle temporal ROI). For all ROIs, we show boxplots that illustrate the distribution of effect-size estimates within each task group. Additionally, we report the results of the meta-analytic mean analysis for our ROIs, i.e., results from the analysis that was described earlier for calculating meta-analytic mean maps. Significance is reported with our default threshold of p < .005, z > 1. For interest, we also report statistical trends found at a more liberal threshold of p < .05, z > 1.

2.2.5. Task group comparisons

To test for differences between the meta-analyses of individual task groups, we carried out linear-models in ES-SDM. Linear models calculate the difference between effect-size estimates from two meta-analyses while taking into account differences in sample size as well as within- and between study variability. This method ensures that differences between groups are not driven by differences in sample size or heterogeneity between the meta-analyses. We performed pairwise comparisons for all 15 combinations of the six task groups in our main analysis. We focus our results on the ROIs which were described in the last section. To account for the multiple comparison problems, we corrected our default threshold of p < .005, z > 1 with the Bonferroni method, giving p < .00033, z > 1. For interest, we also report group differences that are significant only at the uncorrected default threshold p < .005, z > 1.

2.3. Definition of TPJ and mPFC

Two of the most commonly mentioned brain structures in theory of mind research are the temporoparietal junction (TPJ) and the medial prefrontal cortex (mPFC). Both of these labels are only poorly described by classical macroanatomic boundaries and have been used somewhat inconsistently. To characterize and discuss the results from our meta-analysis for the TPJ and the mPFC we rely on recent atlases that map these areas and their subregions based on structural connectivity profiles obtained from diffusion weighted imaging. These atlases are available in fsview (http://fsl.fmrib.ox.ac.uk/fsl/fsview/). Fig. 1 illustrates the connectivity-parcellation clusters of the TPJ and the mPFC. Mars et al. (2012) found that a more anterior connectivity cluster of the TPJ, labeled as TPJa, is interconnected with the inferior frontal gyrus, anterior insula and the SMA. A more posterior connectivity cluster, the TPJp, is interconnected with more dorsal-posterior areas in the IPL, the precuneus, the ventral mPFC, and posterior middle temporal areas. Yet another profile was found for the connectivity cluster in the IPL (dorsal to the TPJa and TPJp), which is, for example, more strongly connected with lateral prefrontal areas. The IPL connectivity cluster can again be subdivided into a number of subclusters along a posterior-to-anterior gradient (Mars et al., 2011). Most relevant for our work are connectivity cluster IPL B, IPL C and IPL D, which lie directly above the TPJ connectivity clusters. Sallet et al. (2013) used connectivity based parcellation
Connectivity based parcellations of TPJ and mPFC

Fig. 1. Illustration of structural connectivity-based mappings of TPJ and mPFC. All maps were thresholded at 25% of participants, so only voxels that belong to any given mask in 25–100% of the participants are labeled. The TPJ was divided in posterior (TPJp), anterior (TPJa), and dorsal (IPL) clusters by Mars et al. (2012). The IPL connectivity cluster can again be subdivided into a number of subclusters (Mars et al., 2011), most relevant for our work are connectivity cluster IPL B, IPL C and IPL D, which lie directly above the TPJ connectivity clusters. For the mPFC, we show connectivity clusters 1–4 derived from a parcellation of the frontal cortex by Sallet et al. (2013).

3. Results

3.1. Individual meta-analyses

3.1.1. False belief vs. photo

Neuroimaging studies have followed developmental psychology in using false belief stories as the prototypical problem for theory of mind reasoning (Fletcher et al., 1995; Gallagher et al., 2000). More recently, it has become apparent that the logical structure of false belief stories systematically differs from that of many control stories. This has been circumvented by introducing the so-called “false” photograph control story (Zaitchik, 1996). In this kind of control condition, subjects are required to represent the outdated content of a physical representation such as a photograph or map. We found 15 imaging studies that used the false belief > photo contrast, which had been first published by Saxe and Kanwisher (2003). Examples of this task are provided in Table 1 in the section ‘False belief vs. photo’.

We performed a meta-analysis on the reported activation maps for the contrast false belief vs. photo stories. Results are shown in the sections ‘False belief vs. photo’ in Fig. 2 and Table 2. We found the largest areas of convergence in left and right temporoparietal cortices. Activation peaks were located in the right posterior superior temporal gyrus corresponding to connectivity cluster TPJp, and the left posterior middle temporal gyrus corresponding to connectivity cluster IPL. Two large areas of convergent activation were also found in the precuneus and in the medial prefrontal cortex. The posterior area included parts of the bilateral precuneus and posterior cingulate gyrus. The anterior area had its activation peak on the border between mPFC connectivity clusters 3 and 4, and extended to more ventral parts of the mPFC and anterior cingulate gyrus. Finally, an additional area of reliable activation was found in the right anterior temporal lobe and adjacent parts of the insula. Jackknife sensitivity analysis showed that all findings of the meta-analysis were highly reproducible, and remained significant for 15/15 combinations of studies using the leave-one-out method. Heterogeneity analysis found significant between study variance only for the peak in the precuneus, whereas all other findings showed no significant variability across studies.

Eleven studies were excluded from the main analyses for the sake of conceptual homogeneity or for methodological reasons. Scholz et al. (2009) was excluded because only ROI data were reported. Another study mixed judgments about false belief and subjective preference (Jenkins and Mitchell, 2010). Other studies relied on various nonverbal formats to study reasoning about false belief, such as comics, pictures and videos. Some of them contrasted activation for false belief versus true belief reasoning (Hooker et al., 2008; Rothmayer et al., 2011; Sommer et al., 2010; Grèzes et al., 2004a, 2004b, 2006), while others contrasted false belief reasoning with reasoning about physical reality (Doehn et al., 2012; Marjoram et al., 2006; Samson et al., 2007). Unfortunately, nonverbal false belief studies were too few in number and too heterogeneous to form a separate task group. In particular, the contrast false belief versus true belief reasoning would deserve a separate task group because of its conceptual uniqueness.

Another 10 studies used methodologically less stringent control conditions and story-based formats to test false belief. We report them in a separate meta-analysis in Supplementary Materials S2. In general, meta-analytic findings for this task category were less robust compared to findings for the false belief vs. photo. One explanation for these weaker findings may be that heterogeneity and methodological variability for these studies was much larger than for the false belief vs. photo. For example, some studies presented relatively long text passages with plenty of information with no relevance for belief-reasoning (e.g., Fletcher et al., 1995). Including these studies in our main analysis would therefore introduce substantial variance which may not be due to clearly identifiable conceptual differences between tasks, but rather due to haphazard methodological differences.

3.1.2. Trait Judgments

Inspired by former findings of brain areas specialized for conceptual knowledge about different classes of inanimate objects (e.g., tools, houses), early studies presenting trait judgment tasks aimed at finding specialized brain areas for conceptual knowledge about persons (e.g., Mitchell et al., 2002). Since then, trait judgment tasks have been very popular in the field. Common to all these studies in this group, the experimental task contains written material which provides information about traits of a person (adjectives, opinions, or personal episodes). The section ‘Trait Judgments’ in Table 1 gives examples. In most of the studies, the described person was not visible; however, three studies presented a photograph of the face of the target person. Different types of control tasks were used in studies of trait judgments. In a number of control tasks attention was diverted away from mental state processing by asking for a lexical judgment on a personality trait word (e.g., Is this word written in upper- or lower-case?). Other control tasks contained words or statements which had no mental-state related content.

We performed a meta-analysis on the reported activations for contrasts of trait-diagnostic information processing > no trait-diagnostic information processing. Sections ‘Trait Judgments’ in Fig. 2 and Table 2 summarize the results. A medial prefrontal area showed absolute highest convergence, with its peak in mPFC.
Table 1
Examples from each task-group in our meta-analysis.

<table>
<thead>
<tr>
<th>Author</th>
<th>Img.</th>
<th>Experimental task</th>
<th>Control task</th>
</tr>
</thead>
<tbody>
<tr>
<td>False belief vs. photo (3 out of 15 studies, for more information see Supplementary Table S1.1) Aichhorn (2009)</td>
<td>fMRI</td>
<td>n = 21 Read a short vignette involving a person holding a false belief. Predict the behavior of that person based on her belief. e.g., ‘Julia sees the ice cream van go to the lake. She doesn’t see that the van turns off to the town hall. Therefore, Julia will look for the ice cream van at the . . .?’ (fake or town hall).</td>
<td>Read a short vignette involving a photograph of the past, and a description how things shown on the photo have changed by now. Answer a question about the outdated scene shown on the photo. e.g., ‘Julia takes a picture of the ice-van in front of the pond. The ice cream van changes to the market place; the picture gets developed. On the picture, the ice-van is at the . . .?’ (pond or market place).</td>
</tr>
<tr>
<td>Saxe (2003)</td>
<td>fMRI</td>
<td>n = 21 Read a short vignette involving a person holding a false belief. Answer a question about her belief. e.g., ‘John told Emily that he had a Porsche. Actually, his car is a Ford. Emily doesn’t know anything about cars so she believed John. When Emily sees John’s car, she thinks it is a . . .?’ (Porsche or Ford).</td>
<td>Read a false-photograph vignette. Answer a question concerning the outdated content in the photo. e.g., ‘A photograph was taken of an apple hanging on a tree branch. The film took half an hour to develop. In the meantime, a strong wind blew the apple to the ground. The developed photograph shows the apple on the . . .? (tree or ground).’</td>
</tr>
<tr>
<td>Lee (2011)</td>
<td>fMRI</td>
<td>n = 26 Read a short vignette involving a person holding a false belief. Answer a question about her belief. e.g., ‘David knows that Ethan is very scared of spiders. Ethan, alone in the attic, sees a shadow move and thinks it is a burglar. David hears Ethan cry for help. David assumes that Ethan thinks he has seen . . .?’ (a spider or a burglar).</td>
<td>Read a false-photograph vignette. Answer a question concerning the outdated content in the photo. e.g., ‘Amy made a drawing of a treehouse three years ago. That was before the storm. We built a new treehouse last summer, but we painted it red instead of blue. The treehouse in Amy’s drawing is . . .?’ (red or blue).</td>
</tr>
<tr>
<td>Trait judgments (3 out of 15 studies, see Supplementary Table S1.2) Ma (2011)</td>
<td>fMRI</td>
<td>n = 30 Read written statements conveying diagnostic information about persons (describing behavior). Then read a single trait-adjective and indicate whether it is consistent with the behavior of that person. e.g., ‘Tolvan gave her sister a hug . . . consistent with “friendly”?’</td>
<td>Read written statement about a person doing something. This behavior is neutral and does not convey trait diagnostic information about the person. Indicate the gender of the person in the sentence. e.g., ‘Tolvan gave her mother a bottle . . . is Tolvan male or female?’</td>
</tr>
<tr>
<td>Mitchell (2002)</td>
<td>fMRI</td>
<td>n = 34 Read an adjective. Indicate whether it can be true for a hypothetical person. e.g., “nervous” . . . can it be true for “David”?’</td>
<td>Read an adjective. Indicate whether it can be true for an object. e.g., “sundried” . . . can it be true for “grape”?’</td>
</tr>
<tr>
<td>Zhu (2007)</td>
<td>fMRI</td>
<td>n = 13 Read a personality trait adjective (e.g., brave, childish) and indicate if it correctly describes a former American president (Bill Clinton).</td>
<td>Read a personality trait adjective (e.g., brave, childish) and indicate if it is written in lower- or upper-case.</td>
</tr>
<tr>
<td>Strategic games (3 out of 9 studies, see Supplementary Table S1.3) Gallagher (2002)</td>
<td>PET</td>
<td>n = 9 Play “stone, paper, scissors” with a human opponent. Select one option (e.g., stone). After that, the option of your opponent is shown. The winner gets rewarded (winner is determined by a set of rules, e.g., stone beats scissors).</td>
<td>Play “stone, paper, scissors” with a computer. You are informed that the computer chooses by a simple algorithm. Select one option.</td>
</tr>
<tr>
<td>Kircher (2009)</td>
<td>fMRI</td>
<td>n = 14 Play the prisoner’s dilemma game (iterated version). You play with a human player for game points. Both players choose a cooperative or defective strategy on each trial. If both players choose defective, they gain almost no game points at all. If both choose cooperative, both gain some game points. If players choose differently, the defective player gains more points.</td>
<td>Play the prisoner’s dilemma game (iterated version). You play with a computer.</td>
</tr>
<tr>
<td>Sripada (2009)</td>
<td>fMRI</td>
<td>n = 26 Play a game like the ultimatum game (iterated version) with a human opponent. Two players must share an amount of money. Player 1 makes an offer how to split the amount, and player 2 can accept (both receive their share) or decline (no one receives money).</td>
<td>Play a game like the ultimatum game (iterated version) with a computer.</td>
</tr>
<tr>
<td>Social animations (3 out of 14 studies, see Supplementary Table S1.4) Blakemore (2003)</td>
<td>fMRI</td>
<td>n = 10 Watch a video animation of two interacting triangles (e.g., surprising one another). Answer questions concerning the contingency (e.g., is there an intention?) between movements of the two shapes.</td>
<td>Watch video animation of two interacting triangles (e.g., surprising one another). Answer questions concerning the physical movement of the first shape (e.g., did the velocity change?).</td>
</tr>
<tr>
<td>Castelli (2000)</td>
<td>PET</td>
<td>n = 6 Watch a video animation of two interacting triangles (e.g., mother and child are playing). Explain verbally what was happening (after fMRI).</td>
<td>Watch video animation of two randomly moving triangles. Explain verbally what was happening (after PET).</td>
</tr>
<tr>
<td>Social animations</td>
<td>fMRI</td>
<td>n = 12 Watch a video animation of simple geometrical shapes depicting a social interaction. Indicate which action was depicted – select a word from several alternatives (e.g., dancing, fishing, sharing . . .).</td>
<td>Watch a video animation of simple geometrical shapes depicting a mechanical action. Indicate which action was depicted – select a word from several alternatives (e.g., billiards, bowling . . .).</td>
</tr>
<tr>
<td>Mind in the eyes (3 out of 10 studies, for more information see Supplementary Table S1.5) Baron-Cohen (1999)</td>
<td>fMRI</td>
<td>n = 12 View photographs of eyes. Indicate which of two words (e.g., concerned versus unconcerned) describes the mental state of that person.</td>
<td>View photographs of eyes. Indicate if the person is male or female.</td>
</tr>
<tr>
<td>Mitchell (2005b)</td>
<td>fMRI</td>
<td>n = 18 View photograph of a face. Indicate how pleased the person was about being pictured.</td>
<td>View Photograph of a face. Indicate how symmetrical the face is.</td>
</tr>
</tbody>
</table>

connectivity cluster 3. The area extended ventrally to mPFC connectivity cluster 4, and dorsal-posteriorly to connectivity cluster 2. Another area of convergent activation was found in the posterior cingulate and parts of the precuneus. Convergence was also found in the temporoparietal cortices, with activation peaks in the right posterior superior temporal gyrus corresponding to connectivity cluster TPp, and the left posterior middle temporal gyrus corresponding to connectivity cluster IPL. In addition, two areas of convergent activation were found in bilateral anterior temporo-parietal lobes. Jackknife sensitivity analysis showed high replicability for bilateral temporo-parietal findings (14/15 combinations left, 15/15 right), medial prefrontal and posterior cingulate findings (15/15 combinations for both peaks). The left anterior temporal findings also showed high replicability (15/15), whereas the right anterior temporal peak could only be reproduced for 11/15 combinations. Heterogeneity analysis found significant between-study variance for all left hemispheric areas (left anterior temporal, left temporo-parietal) and for the posterior cingulate, which indicates that these areas may be activated in some trait judgment tasks more strongly than in others. No significant variance was found for the findings in the medial prefrontal, right temporo-parietal and right anterior temporal cortices.

Twelve studies were excluded from the main analyses of trait judgments. Some studies contrasted activation for other versus self-related judgments (Blackwood et al., 2003; D’Argembeau et al., 2007; Gutchess et al., 2007; Vanderwal et al., 2008) or used another form of mental state judgment as a comparison condition (Cloutier et al., 2011; Ray et al., 2010; Tamir and Mitchell, 2010). Other studies only reported activation for self-related judgments (Schroeter et al., 2010) or reported activation for a mix between self- and other related judgments (Lombardo et al., 2010; Modinos et al., 2009). Pfeifer et al. (2007) was excluded because this study used passive rest as a control condition. Heberlein and Saxe (2005) presented a kind of social animation and asked for a trait judgment. We excluded this task from our analysis because it could not be clearly assigned to one of our task groups (i.e., trait judgments or social animations).

3.1.3. Strategic games

Researchers asked participants in the scanner to play a game with another person. The hypothesis was that feedback from a social partner – indicated by her moves in the game – is spontaneously used to infer her intentions, even if participants are not explicitly told to mindread (Rilling et al., 2004). Then, brain activation measured during strategic games was compared with activation found in studies that explicitly asked for a judgment about mental states, such as false belief. Overlaps in activation were taken to support the hypothesis of an “intentional stance” (e.g., Dennett, 1971), i.e., that people have a disposition to reason about the beliefs, desires and intentions of others to predict behavior (Gallagher et al., 2002). In all strategic games reviewed here, participants were asked to play a game where they could compete or cooperate with another player. Section ‘Strategic Games’ in Table 1 gives examples. Players could not see each other, but they were told about the decision of the other player in the game. The most popular game used by these studies is the classic prisoner’s dilemma. In all experimental tasks, the other player was human. In the control tasks, the other player was a computer (following a simple algorithm).

A meta-analysis was performed on the reported activations for contrasts of playing with a human > playing with a computer. See ‘Strategic Games’ in Fig. 2 and Table 2 for results. The largest area was found in the mPFC with its peak in mPFC connectivity cluster 3. The area extended to the anterior cingulate gyrus ventrally and the right posterior frontal cortex (mPFC connectivity cluster 5) laterally. Another large area was found in the anterior cingulate gyrus and thalamus. Furthermore, we found activation in a right posterior middle temporal area, with subpeaks in connectivity clusters IPL and TPp. A smaller area was also found in the left fusiform gyrus. Jackknife sensitivity analyses showed high replicability for all findings (9/9 combinations) except for the fusiform area, where the peak was only replicated in 6 out of 9 combinations. Heterogeneity analysis found no significant between-study variability for the main findings of the meta-analysis of strategic games.

We excluded 7 studies from the meta-analysis of strategic games. Tomlin et al. (2006) was excluded because they only reported ROI data. Bhatt and Camerer (2005) and Krill and Platek (2012) used control conditions which likely engaged reasoning about mental states. Other studies, using low-level control conditions (Decety et al., 2004; Elliott et al., 2006; Rilling et al., 2008), were excluded for the sake of homogeneity. Another group of studies were excluded because they did not use a subtraction design with activation and control tasks, but computational modeling approaches (e.g., Behrens et al., 2008; Hampton et al., 2008; Yoshida et al., 2010).

3.1.4. Social animations

Social animations were introduced because researchers (Castelli et al., 2000, 2002) sought for a low-level stimulus that provides an alternative to the high-level verbal or cartoon-based materials that predominated the field in the beginning (e.g., Baron-Cohen et al., 1994; Fletcher et al., 1995; Happé et al., 1996). The idea
Meta-analyses for individual task-groups

Fig. 2. Results of separate meta-analyses for task groups. Colors represent probability values from statistical permutation testing (z-values). Maps were thresholded at voxel-wise threshold of p < .005 uncorrected and a cluster extent threshold 10 voxels.

was to investigate whether theory of mind areas found in earlier verbal studies could also be triggered by a very minimalistic and low-level input, which was aimed at showing the existence of a dedicated neurocognitive mechanism for mental-state attribution. Studies in this category presented video animations of simple geometrical shapes, which were introduced by Heider and Simmel (1944). The section ‘Social animations’ in Table 1 gives examples. The majority of studies in this category used the following experimental paradigm: In the experimental task, a movie showed a number of geometrical shapes (e.g., two triangles), moving across the display. The movements portrayed actions which are typical for an intentional or social interaction. In the control condition, a movie showed geometrical shapes in random or purely mechanical movements (e.g., resembling the movement of billiard balls on the table). For each movie, participants were asked to explain/decide if an interaction was portrayed. A few studies in this category showed similar movies in the activation and control tasks, but used different instructions. In the experimental task, participants were asked to focus on the social interactions shown in the movies. In the control task, participants had to focus on the mechanical properties of the geometrical shapes (e.g., indicate if the velocity of the movement had changed).

The meta-analysis for this category was based on contrasts between social or intentional interactions > physical movements. Results are summarized in the sections ‘Social Animations’ in Fig. 2 and Table 2. We found the largest areas of convergence in bilateral temporo-parietal cortices. The right area had its activation peak in connectivity cluster TPJ in the connectivity-based parcellation.
<table>
<thead>
<tr>
<th>Label</th>
<th>Cluster center</th>
<th>z-val.</th>
<th>vx</th>
<th>JK</th>
<th>Het</th>
<th>Individual foci</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>False belief vs. photo (15 studies)</td>
<td>R precuneus (BA 7)</td>
<td>8</td>
<td>−59</td>
<td>35</td>
<td>6.50</td>
<td>1407</td>
</tr>
<tr>
<td></td>
<td>R sup. temporal (BA 39, TPJp)</td>
<td>56</td>
<td>−56</td>
<td>25</td>
<td>13.22</td>
<td>887</td>
</tr>
<tr>
<td></td>
<td>R insula (BA 13)</td>
<td>49</td>
<td>−8</td>
<td>−11</td>
<td>3.36</td>
<td>628</td>
</tr>
<tr>
<td></td>
<td>L mid. temporal (BA 39, IPL D)</td>
<td>−55</td>
<td>−65</td>
<td>27</td>
<td>8.44</td>
<td>960</td>
</tr>
<tr>
<td></td>
<td>L sup. frontal (BA9, cl 3, 4)</td>
<td>−7</td>
<td>58</td>
<td>21</td>
<td>6.29</td>
<td>2095</td>
</tr>
<tr>
<td>Trait judgments (15 studies)</td>
<td>R sup. temporal (BA 39, TPJp)</td>
<td>47</td>
<td>−56</td>
<td>20</td>
<td>2.23</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>R sup. temporal (BA 38)</td>
<td>44</td>
<td>8</td>
<td>−26</td>
<td>2.00</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>R med. front. (BA9, cl 3)</td>
<td>6</td>
<td>51</td>
<td>24</td>
<td>4.38</td>
<td>2276</td>
</tr>
<tr>
<td></td>
<td>L mid. temporal (BA 39, IPL D)</td>
<td>−55</td>
<td>−65</td>
<td>27</td>
<td>2.28</td>
<td>330</td>
</tr>
<tr>
<td></td>
<td>L post cingulate (BA 31)</td>
<td>−3</td>
<td>−56</td>
<td>26</td>
<td>3.64</td>
<td>970</td>
</tr>
<tr>
<td></td>
<td>L mid. temporal (BA 21)</td>
<td>−57</td>
<td>−9</td>
<td>−21</td>
<td>2.82</td>
<td>417</td>
</tr>
<tr>
<td>Strategic Games (9 studies)</td>
<td>R mid. temp. (BA 19, –)</td>
<td>48</td>
<td>−60</td>
<td>16</td>
<td>3.14</td>
<td>422</td>
</tr>
<tr>
<td></td>
<td>L fusiform (BA 19)</td>
<td>−34</td>
<td>−70</td>
<td>−14</td>
<td>2.14</td>
<td>241</td>
</tr>
<tr>
<td></td>
<td>L ant. cing. (–)</td>
<td>−2</td>
<td>2</td>
<td>−4</td>
<td>2.58</td>
<td>657</td>
</tr>
<tr>
<td></td>
<td>L med. front. (BA 32, cl 1)</td>
<td>−2</td>
<td>46</td>
<td>26</td>
<td>4.53</td>
<td>1404</td>
</tr>
<tr>
<td>Social animations (14 studies)</td>
<td>R sup. temporal (BA 13, TPJp)</td>
<td>53</td>
<td>−31</td>
<td>9</td>
<td>4.55</td>
<td>1664</td>
</tr>
<tr>
<td></td>
<td>R thalamus</td>
<td>12</td>
<td>−13</td>
<td>1</td>
<td>2.78</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>R sup. temporal (BA 21)</td>
<td>57</td>
<td>−4</td>
<td>−16</td>
<td>3.06</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>R inf. frontal (BA 45)</td>
<td>57</td>
<td>29</td>
<td>14</td>
<td>2.76</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>R med. frontal (BA 8, cl 3)</td>
<td>6</td>
<td>59</td>
<td>32</td>
<td>2.66</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>L cerebellum</td>
<td>−23</td>
<td>−76</td>
<td>−28</td>
<td>2.79</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>L cerebellum</td>
<td>−31</td>
<td>−45</td>
<td>−20</td>
<td>4.11</td>
<td>704</td>
</tr>
<tr>
<td></td>
<td>L mid. temporal (BA 21)</td>
<td>−62</td>
<td>−15</td>
<td>−13</td>
<td>4.10</td>
<td>1292</td>
</tr>
</tbody>
</table>

Table 2 (Continued)

<table>
<thead>
<tr>
<th>Label</th>
<th>Cluster center</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>vx</th>
<th>JK</th>
<th>Hett</th>
<th>Individual foci</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L mid. frontal (BA 46)</td>
<td>−46</td>
<td>23</td>
<td>19</td>
<td>3.29</td>
<td>297</td>
<td>14</td>
<td>n.s.</td>
<td>L sup./cl 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>L sup./cl 2</td>
<td></td>
</tr>
<tr>
<td>Mind in the eyes (10 studies)</td>
<td>R mid. temporal (BA 19)</td>
<td>56</td>
<td>−63</td>
<td>16</td>
<td>3.31</td>
<td>311</td>
<td>9</td>
<td>n.s.</td>
<td>R mid. temp./p</td>
</tr>
<tr>
<td></td>
<td>R inf. frontal (BA 45)</td>
<td>47</td>
<td>22</td>
<td>6</td>
<td>2.79</td>
<td>28</td>
<td>8</td>
<td>n.s.</td>
<td>R sup. temp./Tpja</td>
</tr>
<tr>
<td></td>
<td>R inf. frontal (BA 9)</td>
<td>60</td>
<td>25</td>
<td>19</td>
<td>2.38</td>
<td>11</td>
<td>5</td>
<td>n.s.</td>
<td>R insula</td>
</tr>
<tr>
<td>L mid. temporal (BA 37)</td>
<td>−51</td>
<td>−62</td>
<td>5</td>
<td>4.84</td>
<td>655</td>
<td>10</td>
<td>n.s.</td>
<td>R mid. temp.</td>
<td></td>
</tr>
<tr>
<td>L sup. cingulate gyrus (BA 24)</td>
<td>−5</td>
<td>8</td>
<td>42</td>
<td>3.75</td>
<td>655</td>
<td>10</td>
<td>n.s.</td>
<td>R med. front.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>L sup. front. (BA 45)</td>
<td>−46</td>
<td>24</td>
<td>7</td>
<td>5.27</td>
<td>1465</td>
<td>10</td>
<td>n.s.</td>
<td>L mid. frontal</td>
</tr>
<tr>
<td>Rational actions (10 studies)</td>
<td>R precuneus (BA 7)</td>
<td>6</td>
<td>−48</td>
<td>38</td>
<td>3.64</td>
<td>894</td>
<td>10</td>
<td>n.s.</td>
<td>L precuneus</td>
</tr>
<tr>
<td></td>
<td>R mid. temporal (BA 21)</td>
<td>60</td>
<td>−45</td>
<td>6</td>
<td>3.57</td>
<td>593</td>
<td>10</td>
<td>n.s.</td>
<td>R mid. temp./p</td>
</tr>
<tr>
<td></td>
<td>R mid. temporal (BA 21)</td>
<td>59</td>
<td>−11</td>
<td>−25</td>
<td>3.54</td>
<td>545</td>
<td>10</td>
<td>n.s.</td>
<td>R mid. temp./p</td>
</tr>
<tr>
<td></td>
<td>L mid. temporal (BA 39, IPL E)</td>
<td>−50</td>
<td>−65</td>
<td>23</td>
<td>3.74</td>
<td>1242</td>
<td>10</td>
<td>n.s.</td>
<td>L sup./Tpja</td>
</tr>
<tr>
<td></td>
<td>L sup. frontal (BA 9, cl 1)</td>
<td>−1</td>
<td>62</td>
<td>23</td>
<td>2.16</td>
<td>31</td>
<td>7</td>
<td>n.s.</td>
<td>L mid. temp./p</td>
</tr>
</tbody>
</table>

Three studies used animations which portrayed only causal relations between mechanically moving objects, but no social content (Blakemore et al., 2001; Fonlupt, 2003; Straube et al., 2011).

3.1.5. Mind in the eyes

The mind in the eyes task was used in fMRI as an attempt to dissociate brain mechanisms subserving general intelligence from those dedicated to social intelligence (Baron-Cohen et al., 1999). This was based on findings that adults with high-functioning autism (Baron-Cohen et al., 1997) as well as parents of children with autism (Baron-Cohen and Hammer, 1997) showed deficits on the mind in the eyes task, but not children with William's syndrome (Tager-Flusberg et al., 1998). Hence, the task was considered as an advanced theory of mind test involving mind-reading (Baron-Cohen et al., 1999). Examples for the task are provided in Table 1. Almost all studies in this category used the following experimental paradigm: In the experimental task, participants saw a photograph of a pair of eyes. In addition, two mental state adjectives were...
shown. Participants were asked to indicate which mental-state
adjective best describes the expression in the eyes on the photo.
The control task also presented a pair of eyes, however, participants
were simply asked to indicate the gender or age of the person dis-
played. A few studies falling into this category showed a complete
face on the photo, or a video of a pair of eyes. Again, a mental state
judgment was compared to a judgment about physical properties
in these studies.

A meta-analysis was performed on the reported activations for
contrasts of mental state judgments × physical judgments (e.g., gen-
der, age), see sections 'Mind in the Eyes' in Fig. 2 and Table 2. A left
inferior frontal area, with a peak BA 45, showed the highest acti-
vation in the map. The area also contained parts of left precentral
and middle frontal gyri, as well as in the left insula. In the right
hemisphere, two smaller lateral prefrontal areas were found in the
inferior frontal gyrus pars triangularis (BA 9 and 45). Further areas
of convergent activation across studies were found in bilateral pos-
terior temporal cortices. The left temporal area was substantially
larger compared to the right. Both areas had their activation peaks
in the posterior middle temporal gyrus, and covered connectivity
clusters TPjp (left and right side) and TPja (right side). The left tem-
poral area further extended to the inferior temporal and fusiform
gyro ventrally. Finally, an area was found in the middle cingulate
gyrus, corresponding to mPFC connectivity clusters 1 and cluster
2 (pre-SMA). Jackknife-sensitivity analysis found perfect replica-
bility for the left temporo-parietal, left middle cingulate and left
inferior frontal findings, and good replicability (9/10 combinations
remained significant) for the right temporo-parietal finding. Lower
replicability was found for the right inferior frontal findings (8/10
and 5/10). Heterogeneity analysis found no significant between-
study variance for the findings of the meta-analysis on mind in the
eyes tasks.

Four other studies found in the literature were not included
in our meta-analysis. Péron et al. (2010) and Baglio et al. (2012)
only reported data for patient populations. Mier et al. (2010) and
Schulte-Rüther et al. (2011) were excluded because they did not
ask for mental states but for basic emotions in their judgments.

3.1.6. Rational actions

The initial idea of studies falling into the category rational
actions was to produce a non-verbal alternative to false belief
stories, which also requires attributing intentions to the protag-
ionist of a story. Another difference is that tasks are not about
false belief but rational action goals. Similar to the idea behind the
social animations task, comic-strips were used to see whether brain
areas dedicated to mental-state attribution responded irrespective
of stimulus input (Brunet et al., 2000). Moreover, the non-verbal
format was used in imaging studies of theory of mind in schizophrenia,
as it circumvents problems linked to speech disorganization in
some patients (Brunet et al., 2003; Sarfati et al., 1997). All tasks in
our meta-analysis had non-verbal material. Examples for tasks are
given in Table 1, section 'Rational actions'. In the experimental
tasks, participants were asked a question about the goal of these
actions, e.g., to predict the likely outcome. The correct answer cor-
responded to the rational (i.e., ordinary) goal pursued by these
simple actions. In the control tasks, questions about non-mental
aspects of the scenes were asked (e.g., about physical causality).

We performed a meta-analysis of the reported activations for
contrasts between inferring action goals × inferring physical
causality. Results can be found in the sections 'Rational actions'
of Fig. 2 and Table 2. Large areas of convergence were found
in bilateral temporo-parietal cortices, with activation peaks in
connectivity clusters IPL, TPjp and TPja on the left side, and
TPja on the right side. Other large areas of convergence were
in the precuneus/posterior cingulate and in right anterior tem-
poral lobe. In addition, we found a small area of convergence
in mPFC connectivity cluster 4. Jackknife sensitivity analysis
showed high replicability for bilateral temporo-parietal, right
anterior temporal and precuneus findings (10/10 combinations
remained significant), and lower replicability for the finding in the
medial prefrontal cortex (7/10). Heterogeneity analysis found no
significant between-study variance for rational actions.

We did not include 5 other studies found in the literature. Two
studies (Brunet et al., 2003; Ciaramidaro et al., 2007) reported data
that had already been used in other studies in our meta-analysis
(so there was a problem of dependent samples). Brune et al. (2011)
only reported data for a patient population. German et al. (2004)
and Saxe et al. (2004a, 2004b) only asked participants to passively
watch human actions, and did not ask about goals or end-states.

3.1.7. Summarizing convergent activations across task groups

The first strategy we used to summarize the meta-analytic
results across task groups was a pooled meta-analysis over all
studies in our sample. This corresponds to the approach taken by
previous voxel-based meta-analyses of theory of mind. Table 3 and
Fig. 3A show results of the pooled meta-analysis. Consistent with
previous work (e.g., Bzdok et al., 2012; Mar, 2011) we found con-
vergent activation along the entire stretch of the bilateral temporal
lobes and parts of the bilateral inferior parietal lobule. Activation
peaks were located in the left and right superior temporal gyri cor-
responding to connectivity clusters TPjp. We found further areas
in the bilateral superior temporal gyri corresponding to connectiv-
ity clusters TPja, and in the bilateral angular gyri corresponding
to connectivity clusters IPL. In addition, we also found large areas
of convergent activation in the precuneus and in the mPFC. The pre-
cuneus showed activation in both hemispheres and had its peak
on the right side. The frontal area had its peak in mPFC connectiv-
ity cluster 3 and extended ventrally to connectivity cluster 4 and
dorsal-posteriorly to connectivity cluster 2/pre-SMA. Furthermore,
relatively small areas of convergence were found in the left inferior
temporal gyrus and in the left fusiform gyrus. Also the right inferior
frontal gyrus showed convergent activation, which was subsumed
under the large right temporal area.

Jackknife sensitivity analysis showed that all of the main find-
ings (L and R TPjp, mPFC connectivity cluster 3, precuneus, left
inferior frontal gyrus) remained significant in 73 out of 73 combina-
tions of studies using the leave-one-out method. Only the peak in
the left fusiform gyrus showed less replicability (remained signif-
ificant in 63 out of 73 analyses). In contrast, heterogeneity analysis
showed significant variance across studies for all of the main find-
ings. High and significant heterogeneity was found in the precuneus
(z = 9.08), left (z = 5.82) and right (z = 8.83) connectivity cluster TPjp,
and in mPFC connectivity cluster 3 (z = 5.29).

Next, we characterized the overlap between task groups by
means of conjunction analysis. As described in the methods section,
we followed three strategies to characterize the overlap between
findings for our task groups. First we calculated a simple conjunc-
tion overlap between significant findings for the six individual task
groups (Fig. 3B). This procedure shows findings which are individ-
ually significant in all six meta-analyses at our default threshold of
p < 0.05 uncorrected, z > 1, and minimum cluster extent 10 voxels.
We found only one cluster of consistent activation in our conjunc-
tion analysis, which was centered in the right posterior middle
temporal gyrus corresponding to connectivity cluster TPjp. This
area comprised 69 voxels and extended to parts of the posterior
superior temporal gyrus.

Our second strategy of calculating the overlap between task
groups was a permutation-based overlap analysis, repeated 10
times for different permutations. Conjunction analysis of all ten
overlay images showed three areas of consistent activation. Results
are reported in Fig. 3C and Table 4. The largest area of conjoint ac-
tivation was found in the right posterior middle temporal gyrus with
Results of pooled meta-analysis.

<table>
<thead>
<tr>
<th>Label</th>
<th>Cluster center</th>
<th>z-val.</th>
<th>vx</th>
<th>JK</th>
<th>Het</th>
<th>Individual foci</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R precuneus (BA 7)</td>
<td>4</td>
<td>-55</td>
<td>34</td>
<td>6.58</td>
<td>171</td>
<td>73</td>
<td>9.08</td>
</tr>
<tr>
<td>R sup. temp. (BA 22, TPJp)</td>
<td>56</td>
<td>-56</td>
<td>18</td>
<td>7.96</td>
<td>2446</td>
<td>73</td>
<td>8.83</td>
</tr>
<tr>
<td>L sup. temp. (BA 22, IPL/TPJp)</td>
<td>-53</td>
<td>-59</td>
<td>20</td>
<td>7.21</td>
<td>1980</td>
<td>73</td>
<td>5.82</td>
</tr>
<tr>
<td>L fusiform (BA 37)</td>
<td>-51</td>
<td>-46</td>
<td>-8</td>
<td>3.36</td>
<td>14</td>
<td>63</td>
<td>n.s.</td>
</tr>
<tr>
<td>L inf. frontal (BA 13)</td>
<td>-44</td>
<td>27</td>
<td>0</td>
<td>4.40</td>
<td>389</td>
<td>73</td>
<td>0.10</td>
</tr>
<tr>
<td>L sup. frontal (BA 9, cl 3)</td>
<td>-1</td>
<td>56</td>
<td>24</td>
<td>7.74</td>
<td>2295</td>
<td>73</td>
<td>5.29</td>
</tr>
</tbody>
</table>

Finally, we followed a third strategy to determine the overlap between task groups. We used the statistically more sensitive approach to determine the overlap between task groups. We used the statistically more sensitive

its peak in connectivity cluster IPL, located at 5 mm distance to connectivity cluster TPJp found for our simple conjunction analysis. On the left side, we found a smaller area of conjoint activation, with its peak in the posterior middle temporal gyrus. In addition, we found an area of conjoint activation in the medial prefrontal cortex, covering parts of mPFC connectivity clusters 3 and 4.

Table 4 Results of conjunction and overlap analyses.

<table>
<thead>
<tr>
<th>#</th>
<th>Label</th>
<th>Cluster center</th>
<th>BA</th>
<th>z-val.</th>
<th>vx</th>
<th>Individual foci</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conjunction (simple overlap) of 6 individual task group meta-analyses</td>
<td>1</td>
<td>R mid. temp./TPJp</td>
<td>49</td>
<td>-56</td>
<td>19</td>
<td>21</td>
<td>2.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>44</td>
<td>-55</td>
<td>18</td>
<td></td>
<td>R mid. temp./TPJp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>56</td>
<td>-55</td>
<td>23</td>
<td></td>
<td>R sup. temp./TPJp</td>
<td></td>
</tr>
<tr>
<td>Conjunction (simple overlap) between two halves of the pooled sample–repeated 10 times with permutation of studies</td>
<td>1</td>
<td>R mid. temp./IPLD</td>
<td>51</td>
<td>-60</td>
<td>20</td>
<td>39</td>
<td>3.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48</td>
<td>-58</td>
<td>30</td>
<td></td>
<td>R mid. temp./IPLD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>64</td>
<td>-46</td>
<td>8</td>
<td></td>
<td>R mid. temp./IPLD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>L mid. temp./TPJp</td>
<td>-55</td>
<td>-59</td>
<td>19</td>
<td>21</td>
<td>3.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-57</td>
<td>-66</td>
<td>25</td>
<td></td>
<td>R sup. temp./TPJp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-57</td>
<td>-66</td>
<td>25</td>
<td></td>
<td>R sup. temp./TPJp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>L med. front./TPJp</td>
<td>-1</td>
<td>54</td>
<td>25</td>
<td>10</td>
<td>3.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-7</td>
<td>60</td>
<td>37</td>
<td></td>
<td>L med. front/cl 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2</td>
<td>61</td>
<td>20</td>
<td></td>
<td>L med. front/cl 3</td>
<td></td>
</tr>
<tr>
<td>Multimodal meta-analysis of 6 task groups</td>
<td>1</td>
<td>R sup. temp./TPJp</td>
<td>62</td>
<td>-58</td>
<td>20</td>
<td>39</td>
<td>2.78</td>
</tr>
<tr>
<td></td>
<td></td>
<td>55</td>
<td>-61</td>
<td>16</td>
<td></td>
<td>R mid. temp./TPJp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>56</td>
<td>-55</td>
<td>27</td>
<td></td>
<td>R sup. temp/TPJp</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>L sup. temp./TPJp</td>
<td>-55</td>
<td>-59</td>
<td>20</td>
<td>22</td>
<td>2.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-52</td>
<td>-47</td>
<td>26</td>
<td></td>
<td>L sup. temp./TPJp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3</td>
<td>44</td>
<td>35</td>
<td></td>
<td>L med. front/cl 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-4</td>
<td>50</td>
<td>39</td>
<td></td>
<td>L med. front/cl 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2</td>
<td>62</td>
<td>24</td>
<td></td>
<td>L med. front/cl 3</td>
<td></td>
</tr>
</tbody>
</table>

Summarizing across task groups

A. Pooled meta-analysis including all studies (n=73)

B. Conjunction (simple overlap) of 6 individual task group meta-analyses

C. Conjunction (simple overlap) between two halves of the pooled sample – repeated 10 times with permutation of studies

D. Multimodal meta-analysis of 6 task groups

Fig. 3. (A) Results of a pooled meta-analysis including all studies from every task group. (B) Results of a conservative conjunction analysis of the six separate meta-analyses. (C) Results of a statistically powerful permutation-based overlap analysis of the six separate meta-analyses. (D) Results of a more liberal multimodal meta-analysis of the six meta-analyses. All results were thresholded at voxel-wise threshold of p < .005 uncorrected and a cluster extent threshold 10 voxels.

multimodal meta-analysis (Radua et al., 2013) to determine areas commonly engaged by the six task groups. Similar to what we found for the permutation-based approach, we found three areas of conjoint activation for the multimodal analysis (see Table 4). Again, the largest area was found in the right temporo-parietal cortex. Its peak was in the right posterior superior temporal gyrus corresponding to connectivity cluster TPJp, at a distance of 13 mm to the peak found for simple conjunction overlap. On the left side, we found a smaller cluster located in connectivity cluster TPJp. Finally, the multimodal analysis found a cluster in the medial prefrontal cortex covering mPFC connectivity clusters 3 and 4.

3.2. Region of interest (ROI) analyses

To characterize task-related differences in brain activation in more detail, we relied on a ROI analysis. We extracted for each study an effect-size estimate (Hedges g) for each ROI. Estimates are displayed as boxplots for separate task groups in Figs. 4–6. Below these boxplots, we indicated with a full circle if activation in that ROI was found significant in the mean analysis, at a threshold of p < .005, z > 1 (determined by permutation tests, see Section 2). Empty circles indicate a trend toward an effect-size above chance level at p < .05, z > 1. Besides boxplots, we also present tables that show the results of pairwise comparisons between task groups for a ROI. Pairwise comparisons were calculated with a linear model that takes into account differences in sample size and variability between meta-analyses. In the tables, full-circles indicate a group-difference at p < .005 Bonferroni corrected, z > 1 (corresponding to an uncorrected p < .00033). Empty circles indicate a trend toward a group-difference at p < .005 uncorrected, z > 1.

3.2.1. Temporo-parietal ROIs

We selected four temporo-parietal ROIs in each hemisphere, based on activation peaks from clusters found in the pooled-meta analysis (indicated in boldface in Table 3). In addition to the left and right temporo-parietal cluster peak, we selected three subpeaks to characterize the extent of this area (for details, see Section 2). Our procedure resulted in bilateral ROIs in the anterior and posterior connectivity clusters of the TPJ: Connectivity-cluster TPJ (left: x = −53, y = −30, z = 10; right: x = 47, y = −35, z = 12) and connectivity cluster TPJ (left: x = −53, y = −59, z = 20; right: x = 56, y = −56, z = 18). In addition, we selected ROIs in connectivity cluster IPL which is distant from the TPJ clusters (left: x = −46, y = −63, z = 41; right: x = 47, y = −61, z = 39). Finally, we selected a posterior middle temporal area, which was located substantially ventral to the TPJ areas (left: x = −66, y = −52, z = 2; right: x = 64, y = −52, z = −3).

As shown in Fig. 4, ROIs #4 and #8 in the left and right posterior middle temporal gyri showed virtually no activations for the task groups false belief vs. photo, strategic games and trait judgments. A common element in these tasks is that they do not show...
Fig. 4. Regions of interest in posterior temporal and parietal brain areas. (a) Box-plots (median; 25th and 75th percentiles; 5th and 95th percentiles) show the distributions of effect-sizes for the studies in each group. Effect-sizes were weighted by intra-study variances. Significant convergence of effect-sizes above zero was determined by randomization tests with SDM; full circles indicate $p < .005$ uncorrected, $z > 1$. Empty circles indicate $p < .05$, $z > 1$. (b) Results of pairwise meta-analytic comparisons between all task groups for the regions of interest. Task groups are abbreviated by capital letters. Pairwise comparisons were calculated by linear contrasts with SDM. Full circles indicate a significant difference between two task groups at $p < .005$, Bonferroni corrected for the number of pairwise comparisons ($n = 15$). Empty circles indicate $p < .005$ uncorrected.
any behavioral cues of agency or intentionality (e.g., movement, facial expression, gaze). For the other task groups, we found activation in these ROIs. In the left posterior middle temporal ROI #4, we found significant activation for social animations and mind in the eyes (z > 2.7), and a trend for rational actions. Pairwise comparisons found that mind in the eyes showed significantly stronger activation than false belief vs. photo and trait judgments (z > 4.3). In the right posterior middle temporal ROI #8, we found significant activation for mind in the eyes and rational actions (z > 2.3) and a tendency for social animations. Pairwise comparisons showed group differences only as tendencies.

ROIs #3 and #7 in connectivity clusters TPJ showed a related pattern to what we found for posterior middle temporal ROIs. Significant activation was found for social animations and rational actions (z > 2.9) and tendentially for mind in the eyes on the right side. On the left side, significant activation was again found for social animations and rational actions (z > 2.0), and tendencies were found for mind in the eyes and false belief vs. photo. Only tendencies and no significant group differences were found in bilateral connectivity clusters TPJ. A strikingly different pattern of activations was found for ROIs #2 and #6 in connectivity clusters TPJ. Significant activation was found for all task groups on the right side (z > 2.3). Moreover, pairwise comparisons showed that activation for false belief vs. photo was higher in this area compared to all other task groups (z > 2.5). On the left side, significant activation was found for all task groups but strategic games (z > 1.9). Again, false belief vs. photo activated the ROI stronger than all other task groups (z > 1.9). In left and right connectivity cluster IPL ROI #1 and #5, significant activation was only found for false belief vs. photo (z > 3.6). In addition, trait judgments activated bilateral connectivity cluster IPL ROIs tendentially. Similar to connectivity cluster TPJ, pairwise comparisons also showed for connectivity cluster IPL ROIs stronger activation for false belief vs. photo compared to a number of other task groups. These were trait judgments, strategic games, mind in the eyes and rational actions on the left side (z > 2.2), and strategic games, social animations, mind in the eyes and rational actions on the right side (z > 1.6).

3.2.2. Cortical midline ROIs

Cortical midline ROIs were selected in a similar fashion to ROIs in the tempo-parietal cortex. We selected the cluster peaks in the medial prefrontal cortex and in the precuneus (see Table 3). In addition, we selected two more local subpeaks in the medial prefrontal cortex, which were (i) at least 20 mm apart from other ROIs and (ii) fell into different connectivity clusters according to the scheme by Sallet et al. (2013). The ROI for the medial prefrontal peak was located in connectivity cluster 4, at x = −1, y = 54, z = 24. More ventrally, a ROI was located at a subpeak within connectivity cluster 3, at x = 3, y = 51, z = −7. Finally, a third medial prefrontal ROI was located substantially more dorsal/posterior in connectivity cluster 2/preSMA, at x = 6, y = 26, z = 55. The ROI in the precuneus was located on the peak at x = 4, y = −55, z = 34. No sub-peaks were found at a distance of more than 20 mm. Effect-sizes of these ROIs are shown in Fig. 5.
ROI #1 in mPFC connectivity cluster 3/ventral mPFC showed significant activation for false belief vs. photo, trait judgments and strategic games ($z > 2.3$). Pairwise comparisons found stronger activation for false belief vs. photo compared to social animations, mind in the eyes and rational actions ($z > 1.8$), as well as stronger activation for trait judgments compared to mind in the eyes ($z = 1.5$). For ROI #2 in mPFC connectivity cluster 4/dorsal mPFC, we again only found significant activation for false belief vs. photo, trait judgments and strategic games ($z > 3.5$). We also found non-significant tendencies for activation for all other task groups. In ROI #3 in mPFC connectivity cluster #2/pre-SMA, we found significant activation for trait judgments ($z = 2.1$) and a trend for mind in the eyes. Group differences were only found in the form of non-significant trends. Finally, in the precuneus ROI #4, we found significant activation for false belief vs. photo, trait judgments and rational actions ($z > 3.0$), as well as a tendency for strategic games.

Pairwise comparisons showed, similar to what we found for the right TPJp ROI, that false belief vs. photo activated the precuneus ROI #4 significantly stronger than all other task groups ($z > 1.6$). In addition, trait judgments activated the precuneus ROI #4 more strongly than mind in the eyes ($z = 2.1$) and tendentially also more strongly than strategic games and social animations.

### 3.2.3. Fronto-temporal ROIs

In accordance with findings from previous literature reviews (Gallagher and Frith, 2003; Mar, 2011), our pooled meta-analysis also found local subpeaks in more anterior parts of the bilateral temporal lobes and in the bilateral inferior frontal gyri (see Table 3). Peaks in the anterior temporal lobes were found near the temporal poles (about 7 mm distance in both hemispheres). Coordinates of the anterior temporal ROIs were $x = -51$, $y = 0$, $z = -19$, and $x = 53$, $y = 0$, $z = -21$, respectively. We also extracted effect-size estimates.
for areas in more central parts of the temporal lobes, located about 20 mm posteriorly to anterior temporal ROIs at \( x = -59, y = -23, z = -8 \), and \( x = 62, y = -22, z = -10 \). For the inferior frontal gyrus, we selected ROIs at \( x = -46, y = 22, z = 8 \), and \( x = 44, y = 20, z = 12 \).

As shown in Fig. 6, ROIs #1 and #4 in the left and right middle temporal gyrus showed activation for social animations and rational actions on the left side \( (z>2.4) \) and for false belief vs. photo and rational actions \( (z>2.0) \) on the right side. In addition, a tendency for activation was also found for social animations on the right side.

Pairwise comparisons found stronger activation for social animations and for rational actions compared to strategic games on the left side \( (z>2.2) \), and stronger activation for false belief vs. photo compared to trait judgments and mind in the eyes on the right side \( (z>2.1) \). For ROIs #2 and #5 in the anterior temporal lobes, we found significant activation for trait judgments, social animations and rational actions on the left side \( (z>2.4) \), and significant activation for false belief vs. photo, social animation and rational actions on the right side \( (z>2.3) \). On the left side, we further found stronger activation for trait judgments compared to strategic games \( (z=1.8) \), whereas no significant group differences were found on the left side. For the left inferior frontal ROIs #3 and #6, we found significant activation for social animations and mind in the eyes \( (z>2.9) \). Pairwise comparisons found stronger activations for mind in the eyes compared to false belief vs. photo, trait judgments, strategic games and rational actions \( (z>3.0) \). For social animations, we found stronger activation compared to false belief vs. photo \( (z=3.1) \), and in trend also compared to other tasks in the left inferior frontal ROI #3. For the right inferior frontal ROI #6, we again found significant activation for mind in the eyes \( (z>2.4) \), and activation trends for trait judgments and social animations. Group differences only emerged as non-significant trends.

4. Discussion

4.1. The role of conceptual variance for meta-analysing theory of mind

Meta-analyses are a useful way to provide objective summaries of the rapidly growing amount of research in the neuroimaging domain. For a large and heterogeneous field such as theory of mind research, it is important to summarize findings with the appropriate level of detail. Different levels can be chosen, and each has its own merits and drawbacks. The most simple and straightforward way to meta-analyse findings is to put together all studies that call themselves a theory of mind task into a pooled analysis. This approach was taken by previous meta-analyses on theory of mind (Decety and Lamm, 2007; Bzdok et al., 2012). A clear benefit of the pooled approach is that it has the best statistical power to detect activations, as it is based on a large sample of all available studies. It will show which activations are robust despite all the variability in the field. Large variability, however, is also a drawback of the pooled approach. A pooled meta-analysis puts together studies using different stimulus materials and control conditions and studies with variable methodological quality. This inevitably will create a lot of noise. To check for variability in our pooled analysis, we performed heterogeneity tests for the main findings. In support of our expectation, we found that all main findings (TPJ, mPFC, precuneus) were accompanied by significant inter-study heterogeneity, which means that the observed between-study variance was larger than that resulting from sampling error alone. The presence of high variability in our pooled analysis indicates that this approach may be missing some details about the functional neuroanatomy of theory of mind. Another potential problem of a pooled analysis is that if some stimulus materials or control conditions have been used more frequently than others, they will have a larger impact on the mean result. Indeed, our literature review shows that some stimulus-materials, like for example false belief vs. photo stories, have been used more often than other stimuli.

A way to address the variability in a meta-analysis is to sort out conceptual or methodological differences between studies, and categorize them along a single stimulus- or task-dimension, as for example verbal versus nonverbal format (Carrington and Baile, 2009), story-based versus non story-based format (Mar, 2011) or enduring versus transient mental state content (Van Overwalle, 2009). The advantage of this approach is that statistical power is still relatively high because the overall sample is split into just two subsamples. However, focusing on a single stimulus- or content-dimension bears the danger of mixing up effects due to other stimulus- or task-aspects not captured by this one dimension, and these may be of equal importance or interest for the overall results. So, single-dimensional meta-analyses may again be pooling data across a variety of different stimulus- and task-aspects, which will bring substantial noise to the results. For example, it could be the case that story-based tasks present a certain mental state (e.g., false-belief) more often than non-story based tasks. Or it could be the case that non-story based tasks present a certain form of control condition (e.g., strategic games: participants are told to interact with a computer) more often than story-based tasks. Such differences would be potential confounds for the comparison between story based and non-story based studies since they have nothing to do with the verbal nature of the tasks. To avoid these problems, it is necessary to separate studies based on the experimental paradigms they used.

4.2. Contribution of the present meta-analysis

For the present meta-analysis, we formed task groups that had comparable stimulus-material, instructions and control conditions. We expected that this approach would reduce the problem of pooling across conceptually and methodologically heterogeneous studies. In support of this assumption, we observed relatively little inter-study heterogeneity in the individual meta-analyses compared to the pooled meta-analysis. No single task-specific meta-analysis showed significant inter-study heterogeneity for peaks in the right TPJ and the mPFC, and only one or two showed between-study variability for peaks in the left TPJ and the precuneus. A potential drawback of our single task group approach is, however, that producing a number of conceptually homogeneous meta-analyses will inevitably produce smaller samples, which reduces statistical power. We therefore used several strategies to check the quality of our results. We used jack-knife sensitivity analyses to evaluate the robustness of each individual analysis. All meta-analyses were recalculated several times, and each time a different study was excluded from the meta-analytic sample. If results remained unchanged in all recalculations of the meta-analysis, this shows robustness against changes of the sample and thus, indicates that meta-analytic findings are replicable. For example, for our smallest task group (strategic games, \( n=9 \)) we found perfect replicability of findings in left and right TPJ and in the precuneus (9 out of 9 leave-one-out recalculations). In contrast, even in the pooled meta-analysis \( (n=73) \) not all findings were perfectly replicable, as the left fusiform activation could be only found for 63 out of 73 leave-one-out recalculations.

Most importantly, we took the problem of small sample size into account when testing for overlaps and differences between task groups. For overlaps, we relied on an alternative approach to check for power-problems due to small sample-size. We not only calculated the simple overlap (conjunction) between individual meta-analyses, as this approach may yield a false impression of selectivity. In addition, we carried out a permutation-based overlap analysis, where we determined the overlap between two halves.
of the pooled-meta analysis sample, each consisting of the studies from three different task groups. This procedure was repeated ten times to create every permutation for assigning six task groups to two sample-halves. Consequently, our procedure identified areas engaged in all task groups, but relied on much larger sample sizes for meta-analyses ($n > 28$). The permutation procedure (see Fig. 1C) turned out to be more sensitive than an analysis of the overlap between the six individual meta-analyses (Fig. 1B). To avoid a false impression of selectivity due to small samples, we will rely on the results from our permutation based procedure in our discussion.

Differences in statistical power must be taken into account when comparing the results of different meta-analyses. For example, when one meta-analysis fails to implicate an area compared to another one, this could reflect differences in sample sizes and study-variability rather than differences between the tasks. To factor this issue in our comparisons, we used a linear model approach which calculates the difference between two meta-analyses while taking into account differences in sample sizes and within- and between study variability. Indeed, our results show a pattern that goes beyond what can be explained by power-differences. For example, when comparing false belief vs. photo tasks with mind in the eyes tasks, we found both areas with overactivation for false belief (e.g., bilateral IPL and TPJ) and areas with overactivation for mind in the eyes (e.g., left IFG). Meta-analytic differences driven by power-differences can be expected only for one direction, i.e., overactivations for the meta-analysis with higher power, but not the other way around.

4.3. Core network for theory of mind

Mar (2011) argued that looking at the overlap between meta-analyses for different paradigm-classes or experiments offers an interesting way for identifying a core-network for theory of mind. The reason is that, whatever is pivotal for theory of mind will be reflected by the convergence of separate methodologies, which differ in idiosyncratic design elements and will not contribute to the overlap. Mar (2011) separated theory of mind research into story based versus non-story based tasks. In the present study, we performed a more detailed separation of task groups. In addition, we performed a permutation based overlap analysis to ensure that our fine-grained study separation does not lead to a power-problem and thus a false impression of selectivity. Our permutation based analysis found three areas of overlap: The mPFC, left TPJ and right TPJ (connectivity clusters TPJ on both sides). The same results were also found with an alternative statistical procedure for determining overlap between meta-analytic maps, a so-called multimodal meta-analysis (Radda et al., 2013). The right TPJ cluster showed the most robust overlap among all findings, as it could even be found in a highly conservative and low-powered conjunction analysis. These results support the conclusion from earlier reviews that a ‘core-network’ for theory of mind exists, i.e., that all sorts of theory of mind tasks consistently engage a particular brain network, including mPFC and bilateral TPJ (e.g., Amodio and Frith, 2006; Frith and Frith, 2006; Mitchell, 2009). This also relates to theories that have proposed a specialized mechanism for mental state attribution, as for example the “theory of mind mechanism” (ToMM, Leslie and Thais, 1992), which was conceived as a mechanism that “kick starts belief and desire attribution” (Leslie et al., 2004). Our results, however, suggest that this mechanism also includes attribution of mental states other than beliefs and desires, as we also found activation in mPFC and bilateral TPJ for tasks like mind in the eyes and trait judgments. The question about functional specificity, i.e., the question whether some parts of mPFC and bilateral TPJ are exclusively (or at least preferentially) engaged in theory of mind, goes beyond the scope of our work. In the following sections, we will discuss the functional roles of the core network’s components as well as the implication of task-related differences that we found.

4.4. Areas for theory of mind

4.4.1. Temporo-parietal junction

The TPJ is roughly characterized as an area at the border between the temporal and parietal lobes surrounding the ends of the Sylvian fissure. It is sometimes also referred to as pSTS, posterior IPL, Ventral Parietal Cortex, Angular Gyrus or Brodmann Area 39. This is linked to the lacking consensus on how coordinates, micro- or macroanatomical landmarks should topographically define that area. The functional heterogeneity of this area has become already evident in patient and brain stimulation studies. On the one hand, damage to the right TPJ frequently leads to an inability to attend to stimuli that are presented in the contralateral visual field (i.e., neglect, see Corbetta and Shulman, 2011) and electrical stimulation of the TPJ results in altered self-awareness. For example, electrical stimulation to the right TPJ leads to the so-called ‘out-of-body experience’ (Blanke et al., 2002), whereas stimulation to the left TPJ leads to the perception of an ‘illusory shadow person’ (Arzy et al., 2006). On the other hand, transcranial magnetic stimulation (TMS) of the right TPJ was found to affect belief reasoning: Young et al. (2010a, 2010b) showed that TMS, participants were for example impaired in taking into account a person’s intention to harm when asked for a moral evaluation of an act of violence. Moreover, two patient studies (Apperly et al., 2004; Samson et al., 2004) reported that damage to the left TPJ leads to selective deficits in false belief reasoning, whereas other cognitive and executive domains remained intact.

In the functional imaging field, studies found that temporo-parietal areas contributing to theory of mind can be functionally divided into a more ventral/anterior part including the STS and a more dorsal/posterior part in the TPJ (e.g., Bahnemann et al., 2010; Gobbini et al., 2007; Perner and Leekam, 2008; Saxe, 2006). However, the only formal literature review looking at task-related differences for theory of mind (Van Overwalle, 2009) found equal activation in the TPJ across a variety of tasks. A potential explanation for the absence of task-related differences is that Van Overwalle (2009) used one large ROI to analyze the TPJ, which did not allow checking for differences between anterior versus posterior or dorsal versus ventral parts. The present study used a whole-brain voxel wise meta-analysis that is capable of detecting such differences. Our results show two things: On the one hand, our conjunction analysis found that all theory of mind tasks activate a part of the TPJ which roughly corresponds to area TPJp. On the other hand, we found task-related differences in our ROI analyses (see Fig. 4): More dorsal/posterior areas in the TPJ (TPJp and IPL) showed stronger activation for false belief vs. photo, and tendentially also for trait judgments, compared to other task. More anterior/ventral areas (TPJp and pMTG) showed stronger activations for social animations, mind in the eyes and rational actions compared to other tasks.

The observed functional dissociation within the TPJ supports previous theories (Gobbini et al., 2007; Perner and Leekam, 2008; Perner and Roessler, 2012). For example, Gobbini et al. (2007) linked ventral/anterior TPJ to processing of overt mental states and dorsal/posterior TPJ to processing of covert mental states. Overt mental states are inherent in perceived actions; covert mental states, such as false beliefs, are not necessarily associated with current actions. Perner and Leekam (2008, see also Perner and Roessler, 2010, 2012) explained the functional separation in the TPJ by drawing on the distinction between teleology and belief-desire reasoning which was found in developmental research. Theory of mind tasks are first of all subserved by teleology, we tell another person’s intention simply by knowing that a certain action is a
rational means to achieve a particular goal. Teleological reasoning (linked to ventral/anterior TPJ) is for example sufficient to understand social animations, mind in the eyes or rational actions. People switch to belief-desire reasoning (linked to dorsal/posterior TPJ) by considering agents’ subjective perspectives of what is needed only when required by the task (perspective tasks). In addition to false belief vs. photo, also trait judgments require awareness of perspective, but for different reasons. Traits are habitual patterns of behavior, thought, and emotion. They are characteristic for a person when the person’s habits deviate from the norm. For instance, the trait word “stubborn” (Murphy et al., 2010) describes a person who, according to Merriam Webster Dictionary, refuses to change her ideas or to stop doing something. In other words, a stubborn person still sees a point in persisting when objectively (from the judging person’s point of view) it is time to give up. In another example, a person is called “anxious” or “nervous” (Mitchell et al., 2002) if she tends to be concerned about situations where one normally has no reason to be anxious, i.e., the person takes a deviant perspective on how dangerous or challenging a situation is. So many traits result from habitually biased perspectives and trait judgments are judgments about whether a person habitually takes a different perspective on certain aspects of life.

The overall pattern found for the TPJ – overlap in the TPJp plus functional differentiations in TPJp and IPL versus TPJa and pMTG – can be linked to a concept of functional specialization referred to as the ‘overarching view’. This concept was proposed by Cabeza et al. (2012) with respect to the IPL. The overarching view assumes that although functional subdivisions within a broad brain region exist, they are graded because each subdivision mediates a particular aspect of a global cognitive function supported by the broad region. In our case, the global cognitive function could be inferring/predicting mental states based on (various forms of) available information about a person. Within that broader region (i.e., the TPJ) sub-regions apply the global function to different types of information which vary according to patterns of functional connectivity.

We looked at whether the previously found structural connectivity networks of TPJa and TPJp (Mars et al., 2012; Bzdok et al., 2013) are reflected in patterns of task-related co-activations of our ROIs. We labeled our results according to a structural connectivity based parcellation (CBP) atlas by Mars et al. (2011, 2012). This atlas is based on white-matter anatomy. Recently, Bzdok et al. (2013) were able to corroborate this parcellation scheme based on two functional connectivity measures, namely resting-state connectivity analysis and meta-analytic co-activation mapping. Consistently, Mars et al. (2012) and Bzdok et al. (2013) reported that a more anterior portion of the TPJ, labeled as TPJa, is interconnected with inferior frontal gyrus, anterior insula and the SMA. A more posterior portion of the TPJ, the TPJp, is interconnected with dorsal IPL, the precuneus, the ventral mPFC, and posterior middle temporal areas.

When looking at similarities between task-related activation patterns of different ROIs, we focus only on areas that relate to the structural connectivity networks of TPJa (IFG) and TPJp (pMTG, IPL, precuneus and mPFC). As summarized in Fig. 7, activation patterns found in ROIs belonging to the same structural connectivity network do not show a perfect correspondence to each other. However, some broad tendencies are observable. TPJa, IFG and pMTG only activate for some of the task groups: social animations, rational actions and mind in the eyes tasks. As discussed above, these tasks present depictions of human action or behavior. On the other hand, TPJa, IFG and pMTG univocally do not activate for false belief vs. photo, trait judgments, and strategic games. Structural connectivity shows a link between TPJa and IFG. However, the pMTG is linked to TPJp rather than TPJa. The cortical midline areas mPFC and precuneus activate most strongly for false belief vs. photo and trait judgments. The IPL activates only for false belief vs. photo (but tendentially also for trait judgments). The latter three areas are all structurally linked to TPJp, which also activates strongly for false belief vs. photo and at an average level for trait judgments. However, the TPJp does not only activate for the latter two task groups, but also for social animations, mind in the eyes and rational actions, which could be due to its connections to the pMTG. Taken together, our ROIs show some support for the idea that the nature and location of activation in the TPJ is, at least in part, reflected in the network of co-activated areas. Original imaging studies are necessary to further clarify the extent to which activation in TPJa versus TPJp can be predicted by connectivity-networks.

4.4.2. Medial prefrontal cortex

Strong activation of the mPFC has already been recognized in early studies on theory of mind, which led to the initial assumption that it is specifically linked to reasoning about belief (Fridh and Fridh, 2001; Gallagher and Fridh, 2003; Leslie et al., 2004). However, this hypothesis has been put into question by more recent patient and imaging data. Bird et al. (2004) reported a case of a patient with an extensive medial prefrontal lesion but no impairment across a wide range of theory of mind tasks, including the ‘strange stories’ test which probes false-belief understanding (see Jolliffe and Baron-Cohen, 1999). Furthermore, while it was found that damage to the left TPJ leads to selective deficits in reasoning about belief, medial frontal damage leads to deficits both in belief reasoning and other cognitive and executive tasks (Apperly et al., 2004; Samson et al., 2004). In a functional imaging study, it was found that the mPFC is equally engaged by stories about persons’ thoughts and by stories about physical appearance or bodily sensations (Saxe and Powell, 2006). Based on this evidence Saxe and Powell (2006) suggested that the mPFC is not specifically engaged in belief-desire reasoning, but more generally involved in processing socially or emotionally relevant information about others (see also Aichhorn et al., 2006). Behrens et al. (2008) showed that the mPFC uses information about others to compute an overall impression of another’s personality. During interactive games, Behrens et al. (2008) found that trial-by-trial changes in activation in the mPFC reflect learning about the ‘social value’ of another player, e.g., the gradual build-up of a prediction about how reliable the other player is. This is consistent with the idea that mPFC is particularly engaged when participants are asked to make judgments about a person’s enduring psychological and social properties, such as for example personality traits (Van Overwalle, 2009). Interestingly, Krienen et al. (2010) has recently shown that mPFC activation during trait judgments is driven by the social closeness of the target-person (e.g., friend or stranger?) rather than by mental or psychological characteristics (e.g., how similar are these traits to my own?).

Although our permutation based conjunction showed that dorsal parts of the mPFC are engaged by all theory of mind tasks, we found in our ROI analyses (see Fig. 5) that activation is particularly strong for false belief vs. photo, strategic games and trait judgment tasks, and less reliable for social animations, mind in the eyes and rational actions. The relevant difference between these tasks may be the time-scale on which socially relevant information is processed. False-belief stories often tell about multiple events in a person’s life, and strategic games present the same opponent across multiple rounds of the game. Trait judgments explicitly ask to think about enduring personality traits. On the other hand, social animations, mind in the eyes and rational actions often show only a single action, behavior or facial expression. Presumably, such information is less relevant for reasoning about a person’s stable personality traits or her ‘social value’.

4.4.3. Precuneus

Little evidence concerning the precuneus exists from patient and brain stimulation research, mainly because of its hidden location within the parietal lobe. Despite strong agreement on the fact
Schematic summary of main results from ROI-based analysis

Fig. 7. Schematic summary of significant activations in ROIs that are linked to the structural connectivity networks of TPJa and TPJp. Labels for ROIs are the same as in previous figures. Task groups are abbreviated by capital letters. Black font indicates that mean-analysis shows significant activation for a task group in that ROI (p < .005 uncorrected, z > 1). Underlined letters indicate that linear-contrasts found stronger activation for this task group compared to others in an area (p < .005, Bonferroni corrected for the number of pairwise comparisons).
that activation of the precuneus is one of the most robust correlates of theory of mind, the area has received relatively little attention in functional accounts. Convergent research shows that the precuneus is implicated in visuo-spatial mental imagery. The precuneus is engaged by motor imagery (e.g., Hanakawa et al., 2003), by mental rotation (Suchan et al., 2002), mental simulation of routes (Ghaem et al., 1997), and by mental model building in deductive reasoning tasks (e.g., Knouff et al., 2003; Kulakova et al., 2013). In addition, the precuneus is engaged in judgments about another person’s visual perspective (Vogele et al., 2004). Based on these findings, Cavanna and Trimble (2006) have proposed that one main function of the precuneus in theory of mind is mental imagery to represent the perspective of another person. Results from our ROI analysis (see Fig. 5) show strong activation for false belief vs. photo, trait-judgments and rational actions. These tasks can all be linked to mental imagery. False belief vs. photo and rational action tasks often tell a fictional story or show a fictional event, which invites mental imaging. Traits are, to some extent, reflected in the habitual patterns of behavior of a person. Therefore, making trait judgments also induces mental imagery of how a person acts or behaves in fictitious situations. On the other hand, social animations and mind in the eyes most clearly do not activate the precuneus, which is consistent with a mental imagery function, because for the latter two tasks the relevant action is shown and does not need to be imagined.

4.4.4. Anterior temporal lobes

Studies of semantic dementia patients (for a review, see Patterson et al., 2007) and transcranial magnetic stimulation (e.g., Lambon Ralph et al., 2009) showed that the anterior temporal lobes are critical for the storage of semantic knowledge. For example, it is assumed that the anterior temporal lobes store semantic concepts which allow generalization of knowledge from one exemplar to another (Lambon Ralph and Patterson, 2008). With respect to social cognition, researchers suggested that the anterior temporal lobes are important for storage and retrieval of semantic scripts (e.g., Frith and Frith, 2003; Gallagher and Frith, 2003). A script refers to general world knowledge which can be seen as a record of which particular activities will take place in a particular setting. An example was given by Gallagher and Frith (2003, p. 77): ‘... the activities associated with the restaurant script would include reading the menu, ordering a drink and getting the bill. If I catch the waiter’s eye and make the gesture of writing on my left palm with my right forefinger he will usually bring me the bill. He correctly interprets my actions on the basis of his knowledge of my likely goals. ...’

Recent research has extended this line of reasoning, and arrived at the conclusion that the anterior temporal lobes provide not only scripts but social semantic concepts in general. For example, Zahn et al. (2007) asked participants to judge the semantic similarity of word pairs which were either descriptive of human social semantic concepts (‘honor’ – ‘brave’) or describing biological function (‘nutritious’ – ‘useful’). Social semantic concepts were specifically associated with an increase in activation in bilateral anterior temporal lobes. Sugiura et al. (2006) showed that the anterior temporal lobes are also important areas for recognition of personally familiar or known persons, which is based on relating semantic information to persons (see also Tsukitani et al., 2010). Ross and Olson (2010) showed that overlapping areas of the anterior temporal lobes are engaged by social animations (comparable to those in our meta-analysis), semantic judgments on social words (comparable to those presented during trait judgments in our meta-analysis), and by stories about persons which require reasoning about intentions. The authors concluded that the anterior temporal lobes contribute to the understanding of implied meaning through access to both general conceptual knowledge and to specific social conceptual knowledge, such as background knowledge about social descriptors (e.g., words like friendly and devious), and knowledge about social rules and social etiquette. Findings from our meta-analysis (see Fig. 6) are in support of this social semantic concept interpretation. Bilateral activation in the anterior temporal lobes was found for trait judgments containing social concept words, as well as for false belief vs. photo social animations and rational actions, which all contain a sequence of actions in a social script (e.g., Gallagher and Frith, 2003; Olson et al., 2007). In addition, some of these task groups also activated more central areas in the middle temporal lobe, whereas activation for trait judgments was restricted to the anterior temporal lobes.

4.4.5. Inferior frontal gyrus

The present meta-analysis found convergent activation in the left inferior frontal gyrus only for the task groups social animations and mind in the eyes. For both types of tasks, evidence has suggested that inferior frontal mirror neuron circuits aid in identification of the observed actions and emotional expressions by a ‘common coding’ mechanism for action and perception (see e.g., Keysers et al., 2010; Keysers and Gazzola, 2007). Functional imaging research shows that the inferior frontal gyrus is engaged in both the execution of actions and in the observation of these actions performed by someone else (for review, see Rizzolatti and Craighero, 2004). Furthermore, Carr et al. (2003) showed that these inferior frontal areas support emotion identification and empathy due to their connections to the limbic system.

The functional significance of the inferior frontal gyrus for processing of abstract displays of implied biological movements was shown in a study by Saygin (2007). The authors found that a group of unilateral stroke patients with lesions to either the left or the right inferior frontal gyri were severely impaired in correctly identifying the human actions depicted by point-light displays (e.g., walking, jogging, ...). With functional imaging, the authors additionally showed that these frontal areas were activated in nonimpaired adults for the identification of actions depicted by point-light displays. With respect to the mind in the eyes task, Shamay-Tsoory et al. (2009) found that a group of patients with lesions in the inferior frontal gyrus had particular difficulties here, but performed unimpaired on a second-order false belief task (Stone et al., 1998). In addition, the patients scored relatively low on affective empathy scales of the Interpersonal Reactivity Index Questionnaire (Davis, 1983), namely the scales ‘emphatic concern’ (e.g., ‘I often have tender, concerned feelings for people less fortunate than me’) and ‘personal distress’ (e.g., ‘being in a tense emotional situation scares me’). Taken together, the patient studies highlight that the inferior frontal gyrus is a crucial structure for particular forms of theory of mind reasoning, which two out of the six task groups from our meta-analysis tap into.

4.5. Limitations

This meta-analysis is intended to give an overview of the rapidly growing field of imaging research on theory of mind. Our focus on standard task-based neuroimaging data omits other evidence which is crucial for understanding the neural underpinnings of theory of mind. First of all, the reviewed imaging data, essentially correlational in nature, should always be supplemented by data related to causality, such as TMS and brain lesion studies. We have incorporated both kinds of evidence in our discussion of brain areas to highlight the importance of causal considerations. Second, our meta-analytic procedure excluded findings from advanced methods of analysis. For example, some studies with strategic games used computational modeling for data analysis (Yoshida et al., 2010; Behrens et al., 2008; Hampton et al., 2008). These studies found that the pSTS computes discrepancies between the behavior that one expects from another player (indicated by the nature of
5. Conclusion

For the present meta-analysis, we formed task groups that had comparable stimulus-material, instructions and control conditions. Meta-analytic assessments of between-study heterogeneity showed that our approach reduces the problem of pooling across conceptually and methodologically different studies. We carried out overlap analyses between task groups, and found that the mPFC and bilateral posterior TPJ (connectivity cluster TPj) showed activation for all theory of mind tasks. This is in line with claims about the existence of a ‘core-network’ for theory of mind, i.e., that all sorts of theory of mind tasks consistently engage a particular brain network (e.g., Amodio and Frith, 2006; Frith and Frith, 2006; Mitchell, 2009). In addition, we performed ROI analysis and found a number of task-related activation differences along with the core-network. For example, we found that more dorsal/posterior parts of the TPJ are particularly engaged for tasks that require processing of mental perspectives (Perner and Leekam, 2008; Perner and Roessler, 2012), namely false belief vs. photo and tendentially trait judgments. More ventral/anterior parts of the TPJ are preferentially activated by tasks which depict rational actions or behavior. Together, this pattern of overlap and surrounding activation differences can be linked to a concept of functional specialization referred to as the ‘overarching view’ (Cabeza et al., 2012), which has been suggested for the IPL. This view assumes that although functional subdivisions within a broad brain region exist, they are graded because each subdivision mediates a particular aspect of a global cognitive function supported by the broad region. In our case, the global cognitive function could be inferring/predicting unobservable mental states based on (various forms of) available information about a person. Sub regions apply this global function to different types of information which vary according to patterns of functional and structural brain connectivity. In support of the connectivity-view, we found in our ROI analyses that similarities between ROI-activation patterns show consistent patterns with respect to the structural brain connectivity of TPJ subdivisions (see Mars et al., 2012).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.neubiorev.2014.01.009.

References


Ma, N.; Vanderkolk, M.; Van Overwalle, F.; Seurinck, R.; Fias, W. 2011. Spon-
taneous and intentional trait infers recruit a common mentalizing network to a different degree: spontaneous infers recruit a larger core area. Social Neuroscience 6, 131–143, http://dx.doi.org/10.1080/17470919.2010.548588.

Malis, G. S.; Lagopoulos, J.; Das, P.; Moss, K.; Berk, M.; Coulston, C. M. 2008. A func-


McAdams, C. J.; Krawczyk, D. C. 2011. Impaired neural processing of social attrib-


Mitchell, J. P.; Cloutier, J.; Banaj, M. R.; Macrae, C. N. 2006. Medial prefrontal disso-


Ross, L. A.; Olson, I. R. 2010. Social cognition and the anterior temporal lobes. Neu-


Samson, D.; Apperly, I. A.; Chiavarini, C.; Humphreys, G. W. 2004. Left temporopari-
tetal junction is necessary for representing someone else’s belief. Nature Neuroscience 7, 499–500.
people thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. Neuroimage 19, 1835–1842.