

Published in final edited form as:

Soc Neurosci. 2009 ; 4(2): 135–152. doi:10.1080/17470910802198510.

Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution

Rajesh K. Kana^{1,3}, Timothy A. Keller¹, Vladimir L. Cherkassky¹, Nancy J. Minshew², and Marcel Adam Just¹

¹ Center for Cognitive Brain Imaging, Carnegie Mellon University

² Departments of Psychiatry and Neurology, University of Pittsburgh School of Medicine

³ Department of Psychology, University of Alabama, Birmingham

Abstract

This study used fMRI to investigate the functioning of the Theory of Mind (ToM) cortical network in autism during the viewing of animations that in some conditions entailed the attribution of a mental state to animated geometric figures. At the cortical level, mentalizing (attribution of mental states) is underpinned by the coordination and integration of the components of the ToM network, which include the medial frontal gyrus, the anterior paracingulate, and the right temporoparietal junction. The pivotal new finding was a functional underconnectivity (a lower degree of synchronization) in autism, especially in the connections between frontal and posterior areas during the attribution of mental states. In addition, the frontal Theory of Mind regions activated less in participants with autism relative to control participants. In the autism group, an independent psychometric assessment of Theory of Mind ability and the activation in the right temporoparietal junction were reliably correlated. The results together provide new evidence for the biological basis of atypical processing of Theory of Mind in autism, implicating the underconnectivity between frontal regions and more posterior areas.

Keywords

autism; mentalizing; Theory of Mind (ToM); mental state; fMRI; underconnectivity

Even though difficulty with Theory of Mind processing in autism has been demonstrated in many studies (e.g., Baron-Cohen et al., 1985; Baron-Cohen et al., 1995; Leekam & Perner, 1991; Perner et al., 1989; Reed & Peterson, 1990; Swettenham, 1996; Swettenham et al., 1996), only a few neuroimaging studies have explored this domain in autism (e.g. Happe et al., 1996; Castelli et al., 2002; see Frith, 2001 for a review). The impairment in Theory of Mind (processing the mental states of others) in autism often results in difficulties with the communications and interactions that underpin everyday life. Attributing mental states (or mentalizing) is perhaps one of the most complex forms of reasoning in which humans engage. Attribution of mental states involves the interplay of a set of subprocesses, such as representation of reality, understanding one's own beliefs and the beliefs of others, and decoupling beliefs from reality. Disruption of such complex information processing is a fundamental impairment in people with autism (Minshew et al., 1997). At the cortical level, such complex information processing requires the integration of the activity of multiple brain regions. According to the underconnectivity theory of autism (Just et al., 2007),

whenever inter-region connectivity and coordination come into play, an underconnected system can manifest impairments, particularly when there is a large load on the system. Theory of Mind processing is supported by a network of several discrete neural structures which require connectivity among themselves to provide the integrated functioning. The purpose of the study reported here was to establish whether the differences in Theory of Mind processing in autism are related to a disruption of processing at the cortical network level and not just at the level of individual cortical areas.

The previously proposed theory of cortical underconnectivity (Just et al., 2004) starts with the propositions that inter-regional connective circuitry in the brain is disrupted in autism, and that patterns of thought that are particularly dependent on integration of frontal and more posterior contributions are disrupted (Just et al., 2007). Furthermore, the theory attributes Theory of Mind deficits and executive function disruption to such underconnectivity. The theory proposes a causal link between the anatomical, physiological (brain activity), and psychological phenomena. Specifically, the theory posits that the communication bandwidth among cortical areas, particularly between frontal and posterior areas, is lower in autism than in typical participants. Bandwidth, which refers here to the amount of information that can be transmitted between cortical centers per unit time, is a critical factor in the performance of a computational network. Brain imaging research has been very clear in showing that human thought involves a network of cortical areas whose activity is coordinated (synchronized), and that coordination has to be based on inter-regional communication, using the white matter tracts that provide the anatomical connectivity. The fMRI findings emerging in the last few years indicate that the synchronization between frontal and posterior areas is lower in autism (Just et al., 2006; Just et al., 2004; Kana et al., 2006; Koshino et al., 2005; Villalobos et al., 2005). For example, in a Tower of London (TOL) problem-solving task, which entails activation of both frontal and parietal areas, the synchronization between the frontal and parietal areas is lower in autism than in a control group.

Neuroimaging studies of Theory of Mind processing have identified a frontal-posterior network of brain regions that include the medial prefrontal cortex, the temporal-parietal junction at the top of superior temporal sulcus, and the temporal poles adjacent to the amygdala (Fletcher et al., 1995; Gallagher et al., 2000; Vogeley et al., 2001; Castelli et al., 2000, 2002; Brunet et al., 2000). There has been some debate concerning the differential roles of frontal versus posterior regions in Theory of Mind processing. Although the study reported below does not focus on differentiating the contributions of various brain areas to Theory of Mind processing, it is nevertheless useful to acknowledge that different areas are very likely playing different roles. Some researchers argue that while the medial frontal cortex is directly involved in mental state reasoning and decisions, the right temporoparietal junction collates the cues that are the input to mentalizing (Gallagher & Frith, 2003). On the other hand, others argue that the right temporoparietal junction is the central player in mental state reasoning, not only collating but also processing the cues to compute a mental state (Saxe & Kanwisher, 2003). Regardless of the precise computation that each area performs in ToM processing, it appears that the two regions function in concert with each other. It is this coordination among the contributing areas that is the focus of this article.

The communications with the frontal components of the ToM network are of special interest here, because of the hypothesis of lower frontal-posterior functional connectivity in autism. A number of studies indicate that the role of frontal lobe regions in processing Theory of Mind is major. Evidence for this view comes in large part from lesion research. Patients with damage to frontal cortex are frequently impaired in the performance of Theory of Mind tasks (Stuss et al., 2001). Damage to the anterior prefrontal regions has been associated with impaired self-awareness for the appropriateness of social interactions, judgment and

planning difficulties (Stuss, 1991), as well as impaired awareness of the mental states of others (Stone et al., 1998; Stuss et al., 2001).

It may not be just the frontal activation that is critical, but the synchronization of frontal activity with activity in posterior Theory of Mind areas, as well as other areas, which may be important in accomplishing mentalizing. From an anatomical perspective, major afferents to the medial frontal regions come from the dorsolateral prefrontal cortex, temporal pole, anterior superior temporal gyrus, parietotemporal cortex, and posterior cingulate cortex (Barbas et al., 1999; Carmichael & Price, 1996; see Amodio & Frith, 2006 for a review). Therefore, the disruption of connections from and to the frontal lobe may affect higher cognitive functions, such as Theory of Mind processing.

The present study investigates the synchronization of brain activation in autism using a Theory of Mind task involving an animation of two interacting triangles. Theory of Mind processing in autism using these animations has been studied previously using PET (Castelli et al., 2002). The rationale for doing the present study is that: 1) the current study used the higher temporal resolution of fMRI to examine the synchronization of brain activation (functional connectivity) in autism in a Theory of Mind task for the first time, and 2) the current task targeted the explicit online assessment of mental state attribution in people with autism. We predicted reduced brain activation in autism in Theory of Mind processing regions such as medial frontal cortex and right temporoparietal junction. This hypothesis is based on the findings from studies of impairment in Theory of Mind (Castelli et al., 2002), and perception of biological motion (Blake et al., 2003) in autism. The novel portion is our prediction of functional underconnectivity in autism during mentalizing, especially in connections involving the frontal lobe. This prediction is based on previous findings of functional underconnectivity in autism in a wide variety of higher cognitive tasks in autism (Just et al., 2006; Kana et al., 2006; Castelli et al., 2002). The underconnectivity prediction is also based on evidence from structural brain imaging studies which indicate abnormalities in autism of the white matter that provides the communication channels between brain areas (Herbert et al., 2003, 2004; Keller et al., 2007; Barnea-Goraly et al., 2004).

Method

Participants

Twelve high-functioning adult individuals with autism (mean age 24.6 years) and twelve typical control participants (mean age 24.4 years) were included in the analyses (Full Scale and Verbal IQ scores of 75 or above). Participants were matched on the basis of age and IQ (see Table 1). Among the 12 participants in the autism group, two were females and one participant was left-handed. In the control group, all participants were males and one was left-handed. Data from eight other participants (5 participants with autism and 3 controls) were discarded due to excessive head motion. The diagnosis of autism was established using two structured research diagnostic instruments, the ADI-R (Autism Diagnostic Interview-Revised, Lord et al., 1994) and the ADOS-G (Autism Diagnostic Observation Schedule-Generic, Lord et al., 2000), supplemented with confirmation by expert opinion. None of the participants were diagnosed with Asperger syndrome. Potential participants with autism were excluded on the basis of an associated infectious or genetic disorder, such as fragile-X syndrome or tuberous sclerosis. Potential control participants and participants with autism were also excluded if found to have evidence of birth asphyxia, head injury, or a seizure disorder. Exclusionary criteria were based on neurologic history and examination, and chromosomal analysis.

The control participants were medically healthy community volunteers recruited to match the participants with autism on age, Full Scale IQ, gender, race, and family of origin

socioeconomic status, as measured by the Hollingshead method. Potential control participants were screened by questionnaire, telephone, face-to-face interview, and observation during screening psychometric testing to determine eligibility. Exclusionary criteria, evaluated through these procedures, included current or past psychiatric and neurologic disorders, birth injury, developmental delay, school problems, acquired brain injury, learning disabilities, and medical disorders with implications for the central nervous system or those requiring regular medication. Potential control participants were also screened to exclude those with medical illnesses or a family history of autism, developmental cognitive disorder, affective disorder, anxiety disorder, schizophrenia, obsessive compulsive disorder, or other neurologic or psychiatric disorders thought to have a genetic component (in first-degree relatives or self), medications that affect the CNS, hypertension, diabetes, substance abuse (self, or first-degree relative), steroid use (extreme use such as steroids used in inhalers for asthmatics) and autism in first-, second- or third-degree relatives. Each participant signed an informed consent that had been approved by the University of Pittsburgh and Carnegie Mellon University Institutional Review Boards.

Experimental paradigm

This experiment compared autism and control participants on attributing mental states to the movement of geometrical figures. The animations were generously provided by Dr. Fulvia Castelli, who performed the original PET study of this task (Castelli et al., 2002). All animations featured two “characters”, a large red triangle and a small blue triangle, moving about on a framed white background.

Three types of animations were used in this study: the Theory of Mind (ToM), the Goal-Directed (GD), and the Random (RD), formed three experimental conditions. In the ToM condition, the geometrical figures engaged each other with intentional action/interaction involving thoughts and feelings. For example, a character’s actions were determined by what the other character thought. In the GD condition, the geometrical figures engaged in an interaction with each other in a simple purposeful level. For example, one character’s actions were determined by what the other character did. In the RD condition, the geometrical figures did not engage with each other at all. For example, all character actions were random and unassimilated. There were three trials each of ToM, GD, and RD stimuli, each trial containing one animation. The basic visual characteristics depicted in the three types of animations were similar in terms of shape, overall speed of motion, orientation changes, and total duration (between 26 and 47 sec).

After the presentation of each animation, four words were presented on the screen, and participants were asked to make a forced-choice judgment about which of the words best described the action depicted in each animation. (Note the dissimilarities to Castelli et al.’s (2002) procedure, which required only passive viewing of the animations and reporting after the scan what had happened in the animation). The correct response was always an accurate description of the animation (determined as the most frequently generated description in a norming study). Another response alternative was an inaccurate description of the animation, but of the appropriate category (e.g., in the ToM condition, this incorrect response referred to a mental state, but not the correct mental state). The other two answers were inaccurate descriptions of the animation that could have applied to animations in the other two conditions (i.e., in the ToM condition, these incorrect response alternatives referred to a random motion and a goal-directed motion, in the GD condition they referred to a theory of mind motion and a random motion, and in the RD condition they referred to a theory of mind motion and a goal-directed motion). For example, for a ToM animation that depicted “coaxing”, the foils were pushing (GD), surprising (ToM), and spinning (RD). Participants made their responses using two two-button mice, one held in each hand. Each button corresponded to one of the four multiple-choice answers. Responses were accepted

for 15s from the end of each animation. The presentation of each animation constituted a separate event in the experimental design. The animations were presented in blocks of 3, one from each condition, with a separation of 6 sec between trials within a block. The onset of each animation was synchronized with the beginning of a TR. A 30-second fixation condition was presented between each block to provide a baseline measure of brain activation with which to compare each experimental condition. In this fixation condition, participants fixated on a centered asterisk without performing any task. A practice animation belonging to the RD condition was presented at the beginning of the experiment to familiarize the participant with the task.

We also administered a separate neuropsychological test outside the scanner to get a different measure of participants' ToM abilities. The Happe Strange Stories Test for the assessment of theory of mind was administered as a neuropsychological test just to get an independent out-of-scanner measure of the Theory of Mind abilities of the participants with autism.

Prior to testing in the scanner, participants were familiarized with the task using one example of each of the three conditions that was not presented again during the fMRI study. Familiarization also entailed the use of a scanner simulator to acclimate the participants and to attain motion quality standards.

fMRI procedures

All imaging data were acquired at the Brain Imaging Research Center (BIRC) co-administered by Carnegie Mellon University and the University of Pittsburgh on a 3-Tesla Siemens Allegra scanner. The stimuli were rear projected onto a semi-translucent plastic screen and participants viewed the screen through a mirror attached to the head coil. For the functional imaging, a gradient echo, echo-planar pulse sequence was used with TR=1000 ms, TE=30 ms, and a flip angle of 60°. Sixteen adjacent oblique-axial slices were acquired in an interleaved sequence, with 5-mm slice thickness, 1-mm slice gap, a 20 × 20 cm FOV, and a 64 × 64 matrix, resulting in in-plane resolution of 3.125 × 3.125 mm. A 160-slice 3D MPRAGE volume scan with TR=200 ms, TE=3.34 ms, flip angle=7, FOV=25.6 cm, 256 × 256 matrix size and 1-mm slice thickness, was acquired at the same orientation as the oblique-axial functional images for each participant. This structural scan was used for making measurements of corpus callosum size.

Distribution of activation

To compare the participating groups in terms of the distribution of activation, the data were analyzed using SPM99. Images were corrected for slice acquisition timing and head motion, and they were normalized to the Montreal Neurological Institute (MNI) template, resampled to 2 × 2 × 2 mm voxels, and smoothed with an 8-mm Gaussian kernel to decrease spatial noise. High-pass filtering and global scaling were performed on each participant's data. Statistical analysis was performed on individual and group data by using the general linear model as implemented in SPM99 (Friston et al., 1995). Group analyses were performed using a random-effects model. Contrasts reflecting the group differences in the distribution of activation relative to fixation, and group-by-condition interactions were computed. For the group difference contrasts, possible differences in deactivation (relative to fixation condition) were excluded because previous studies have shown that medial prefrontal areas that activate in response to Theory of Mind tasks also show deactivation relative to a resting baseline for a variety of other cognitive tasks (e.g., Cherkassky et al., 2006). This masking procedure ensured that group differences were not due to greater activation during fixation in one of the groups relative to the other. An uncorrected height threshold of $p < .005$ and an extent threshold of 6 voxels were used.

Functional connectivity

The functional connectivity was computed (separately for each participant) as a correlation between the average time course of signal intensity (The average time course of signal intensity for an ROI was calculated separately within each participant by averaging the signal intensity separately for each time point (image) across activated voxels within the ROI) of all the activated voxels in each member of a pair of regions of interest (ROIs). Fifteen functional ROIs were defined to encompass the main clusters of activation in the group activation map for each group in the ToM-Fixation contrast. Labels for these 15 ROIs [the superior medial frontal gyrus (SMedFG), left and right inferior orbital frontal cortex (LIOFC and RIOFC), left and right precentral gyrus (LPrecen), right inferior frontal gyrus (RIFG), left and right inferior parietal lobes (LIPL and RIPL), left and right superior parietal lobes (LSPL and RSPL), right middle temporal gyrus (RMTG), right superior temporal gyrus (RSTG), left and right fusiform gyrus (LFFG and RFFG), and left and right inferior occipital gyrus (LIOG and RIOG)] were assigned with reference to the parcellation of the Montreal Neurological Institute (MNI) single-subject T1-weighted dataset carried out by Tzourio-Mazoyer and colleagues (Tzourio-Mazoyer et al., 2002). For each of these ROIs, a sphere was defined for each cluster (with a radius ranging from 4 to 8 mm) that best captured the cluster of activation in the map for each group. The ROIs used in the analysis were each the union of the two spheres defined for the two groups (A complete table of the locations and sizes of each of the ROIs within each group can be found in the supplementary material available online). Within these ROIs data were extracted from voxels that exceeded a t -threshold of 4.5 (corresponding approximately to a significance level of $p < .10$, corrected for multiple comparisons) in at least one contrast between an animation condition and the fixation baseline from the individual participant's first-level analysis. In addition, we required that at least 12 voxels were active within a ROI for an individual participant to ensure a stable estimate of the time-course. The activation time course extracted for each participant over the activated voxels within the ROI originated from the normalized and smoothed images that were high-pass filtered and had the linear trend removed. The correlation was computed on the images belonging to the random, goal-directed, and Theory of Mind conditions, so it reflects the interaction between the activation in two areas while the participant is performing the task and not during the baseline condition. Fisher's r to z transformation was applied to the correlation coefficients for each participant prior to averaging and statistical comparison of the two groups.

Functional connectivity was measured for each participant in each group for the ToM, GD, and RD conditions using the 15 functional ROIs described above. On the basis of the prediction that frontal-posterior networks would show significant underconnectivity in autism (Just et al., 2007; Kana et al., 2006), these 15 ROIs were grouped into larger regions on the basis of lobe (frontal, parietal, temporal, or occipital), and then functional connectivity measures for these groups of ROIs were obtained for each participant by averaging the connectivities of all of the relevant ROI pairs. This resulted in 10 "networks" for which connectivities were aggregated: six inter-lobe connectivities (frontal-parietal, frontal-temporal, frontal-occipital, temporal-parietal, temporal-occipital, and parietal-occipital), and four intra-lobe connectivities (within frontal, temporal, parietal, and occipital). In addition, we calculated functional connectivity between the frontal centers of the Theory of Mind network (medial frontal, orbitofrontal) and the posterior centers (right middle and superior temporal gyri, temporoparietal junction).

Corpus callosum morphometry

The cross-sectional area of the midsagittal slice of the corpus callosum was measured using the parcellation scheme described by Witelson (1989). The seven regions of the corpus callosum defined in this scheme include the rostrum, genu, rostral body, anterior midbody,

posterior midbody, isthmus, and splenium. The corpus callosum size was normalized (divided by) by the total grey and white matter volume for each participant. The grey matter, white matter and CSF volumes were measured for each subject by segmenting the T1-weighted structural brain image into three masks. The segmentation was performed by SPM2 routines. The outer contour of the corpus callosum was manually traced (with an inter-rater reliability of .87), and then interior segmentation, area, and length computations were performed by image processing software.

Results

Overview

The central findings indicate that during the attribution of mental states to animated figures, participants with autism (relative to control participants): 1) showed reduced functional connectivity within the Theory of Mind network as well as in the connections between the frontal lobe and more posterior regions (and to our knowledge, this is the first finding linking Theory of Mind processing in autism to the lowered synchronization of activation across relevant brain areas); 2) showed reduced levels of brain activation in the frontal components of the Theory of Mind network (medial prefrontal cortex, anterior cingulate cortex, and orbitofrontal gyrus); 3) showed a positive correlation between Theory of Mind ability (measured psychometrically) and the activation in right superior temporal gyrus; and 4) the midsagittal area of the rostrum of the corpus callosum was found to be smaller in autism than in controls. As described in more detail below, the findings are consistent with disrupted ToM processing in autism being underpinned by frontal-posterior functional underconnectivity.

Behavioral Results

The behavioral results reported here are based on performance in the fMRI task. The autism and control groups performed similarly. There was no statistically reliable difference between the autism and control groups in either reaction time [$F(1,22) = 1.20, p = 0.29$] or error rate [$F(1,22) = 2.50, p = 0.13$], nor was there a statistically reliable interaction between the groups and conditions for either response measure [$F(2,44) = 0.82, p = 0.45$] and [$F(2,44) = 1.85, p = 0.17$]. There was an effect of condition overall, such that the Theory of Mind condition showed significantly longer reaction times [$F(2, 44) = 25.12, p < 0.001$] and significantly higher error rates [$F(2,44) = 14.63, p < 0.001$].

Contrast between ToM and random motion for both groups

In the ToM condition, both the autism and the control group showed activation in regions associated with Theory of Mind, such as the medial frontal gyrus, orbitofrontal gyrus and right middle temporal gyrus. However, the autism group clearly showed reliably less activation in medial frontal regions compared to control participants (see Figure 1 and Table 2), as further described in the next section.

Group differences in brain activation

There were statistically reliable group differences in the contrast between the Theory of Mind condition and the random animations. The autism group showed reliably lower brain activation compared to controls in several regions in the frontal cortex, including the medial frontal gyrus, anterior paracingulate cortex, anterior cingulate gyrus, and inferior orbital frontal gyrus, as shown in Figure 2. All of these regions have been found to co-activate in a “social” processing brain network in typical populations (Adolphs, 2003; Winston et al., 2002; Blakemore, Winston & Frith, 2004; Brothers, 1990; Castelli et al., 2002; Gallagher & Frith, 2003).

In contrast to the lower activation shown by the autism group in frontal parts of the ToM network, there was no statistically reliable group difference in activation in the right posterior superior temporal sulcus region itself (a key posterior Theory of Mind area centered at $x = 52$, $y = -46$, $z = 24$; see Castelli et al., 2002). However, the autism group showed more activation than the control group in a region slightly more anterior to the superior temporal sulcus ($t(22) = 3.41$, $p = 0.001$, centered at $x = 56$, $y = -28$, $z = 18$). Overall, the autism group showed lower activation in frontal Theory of Mind regions while showing no difference in activation in the posterior Theory of Mind region.

Functional connectivity

The key hypothesis regarding functional connectivity was that in the ToM condition, synchronization would be reduced in the group with autism relative to controls between the frontal Theory of Mind areas (medial frontal, orbitofrontal) and posterior ToM areas (right middle and superior temporal gyri, temporoparietal junction). A planned contrast between the two groups comparing this frontal-posterior within-ToM-network functional connectivity within the ToM condition confirmed this prediction [autism mean = 0.20; control mean = 0.48; $t(19) = 2.65$, $p < 0.05$]. Note that this result is not due to a failure of the group with autism to activate either frontal or right temporoparietal areas in the ToM task. Both groups showed activation in these regions for the contrast between ToM and the fixation baseline, and this contrast was used to define the ROIs for extracting the time course of signal intensity (see Supplementary Table). In addition, the procedure for extracting the time-course of signal intensity from an individual participant included only voxels found to be activated for the participant in at least one of the contrasts of an animation condition with this baseline condition. This procedure eliminated consideration of functional connectivity in the frontal-posterior ToM network for three of the participants with autism (all normal controls had activation within the defined ROIs), but prevents differences in the response of frontal areas to ToM task from biasing the group estimate of functional connectivity by taking into account data from only those participants who activated the regions. Note also that functional connectivity measures the degree of synchronization of signal intensity among regions across time rather than the amplitude of the response during the presentation of the animation.

Because reduced frontal-posterior functional connectivity in autism has previously been established in a variety of cognitive tasks and even in a resting baseline, we also tested the specificity of the difference to the requirement for ToM processing by examining the functional connectivity in the RD and GD conditions. Participants with autism showed lower functional connectivity than controls in all three conditions, although this group difference was smallest in the GD condition, as shown in Figure 3. A 3 (Condition) \times 2 (Group) mixed ANOVA revealed a reliable main effect of Group [$F(1, 19) = 7.56$, $p < .05$], indicating reduced overall functional connectivity in the group with autism, and a reliable main effect of condition [$F(2, 38) = 7.75$, $p < .005$], indicating that across groups the functional connectivity was modulated by the type of animation. There was a trend toward a Condition \times Group interaction [$F(2, 38) = 2.27$, $p = .12$]. Tests of the simple main effects of group within each condition indicated that in addition to expected reduction in functional connectivity in the ToM condition, there was also a reliable group difference in the RD condition [$t(19) = 2.58$, $p < .05$], but not in the GD condition [$t(19) = 1.49$, $p = .15$]. Within the autism group there was only a marginal effect of Condition on functional connectivity in the frontal-temporal ToM network [$F(2, 16) = 3.63$, $p = .05$], resulting from lower connectivity in the RD condition than in either the GD condition [$t(16) = 2.51$, $p < .05$] or the ToM condition [$t(16) = 2.11$, $p = .05$]. In contrast, within the control group there was a reliable Condition effect [$F(2, 22) = 6.82$, $p < .005$], with significantly greater functional connectivity in the ToM condition than either the RD condition [$t(22) = 3.56$, $p < .05$] or the

GD condition [$t(22) = 2.64, p < .05$]. Thus, the overall pattern shown in Figure 3 suggests that while normal control participants show increased functional connectivity between frontal and right posterior temporal-parietal areas only when the task demanded Theory of Mind, the group with autism showed increased functional connectivity in this network for both the Goal-directed and Theory of Mind conditions, but did not approach the degree of synchronization of activation between regions in the Theory of Mind condition as control participants.

In order to fully characterize the functional connectivity patterns involving all of the regions with substantial activation in the ToM condition, an ANOVA of the mean inter-lobe and intra-lobe connectivities compared the mean functional connectivities of the two groups averaged within each group over the pairs of functional ROIs that were in different cortical lobes. Thus groups were compared across 10 possible pairings (networks) of the 4 lobes (6 inter-lobe networks and 4 intra-lobe networks): frontal-parietal, frontal-temporal, frontal-occipital, temporal-parietal, temporal-occipital, parietal-occipital, frontal-frontal, temporal-temporal, parietal-parietal, and occipital-occipital. This 2 (Group) \times 10 (Network) mixed ANOVA revealed a reliable main effect of network [$F(9, 153) = 26.46, p < 0.0001$], a reliable main effect of Group [$F(1, 17) = 6.76, p < 0.05$], and no reliable Network \times Group interaction, indicating that even in a task targeting Theory of Mind processing, widespread reductions in functional connectivity are found in autism. Means and simple main effects of Group are shown in Table 3. As the table indicates, the group with autism showed reliably lower functional connectivity between frontal regions and posterior regions (temporal, parietal and occipital) compared to control participants. In addition, reliably lower functional connectivity was found in the autism group in the functional connectivities between temporal and occipital areas. In the intra-lobe networks, the autism participants showed reliably lower connectivity in parietal-parietal as well as occipital-occipital connections.

Correlation between brain activation and mentalizing ability

The brain activation in two regions of interest (the medial frontal gyrus, and the right posterior superior temporal gyrus), which are associated with processing of Theory of Mind, was correlated with the Theory of Mind score of the autism participants from the Happe strange stories test (Happe, 1994). The activation measure used in the correlation was the contrast of parameter estimates (beta-values) from the general linear model extracted from individual participants (anatomical region of interest) for the Theory of Mind vs. Random contrast. The Theory of Mind score was reliably positively correlated [$r = 0.57, t(10) = 2.21, p < 0.05$] with the activation in right posterior superior temporal gyrus (see Figure 3). Although the correlation was positive ($r = 0.26$) for the medial frontal gyrus ROI, the effect was not statistically significant. (The Theory of Mind scores were not available for the control participants, but one might expect scores near ceiling without much variation).

Corpus callosum size

To investigate a possible anatomical basis of frontal-posterior functional underconnectivity and frontal underactivation among the group with autism, we compared the size of the corpus callosum in anterior regions (the rostrum and genu) between the groups. The results showed that the rostrum was reliably smaller in participants with autism compared to control participants [$t(22) = 2.547, p < 0.02$], although no such effect was found for the genu. The rostrum of the corpus callosum connects orbital prefrontal regions in the two hemispheres, regions previously found to be involved in Theory of Mind processing (Sabbagh, 2004; Stone et al., 1998). Moreover, the present study found reduced activation in a left orbital prefrontal region in the group with autism relative to controls, providing a link between lowered activation in ToM in autism and connective white matter. Unlike several previously reported studies (Just et al., 2006; Kana et al., 2006; Cherkassky et al., 2006), this analysis

found no significant correlation between the size of any corpus callosum segment and the functional connectivities between the areas that the segment connects.

Discussion

The most important finding of the current study is that the participants with autism showed functional underconnectivity among several brain areas involved in Theory of Mind processing. As predicted, the frontal Theory of Mind regions (medial frontal gyrus, anterior paracingulate, orbital frontal gyrus) were found to be functionally underconnected with the posterior Theory of Mind regions (right middle and superior temporal gyrus) in autism during mentalizing. In addition, underconnectivity was found in the connections between the frontal lobe and other more posterior regions, and within the occipital and parietal lobes. Castelli et al. (2002) reported a correlation across autistic and control participants between the amount of activation (averaged over the entire course of the task) in occipital and temporal regions in a PET study. The current study, using the same stimuli, provides new evidence of a lower degree of synchronization (with observations obtained once per second) between frontal and posterior areas in autism during the viewing of these animations. The fMRI measures provide critical detail that speaks to the underlying mechanisms that distinguish the two groups, namely the lower degree of synchronization and hence coordination among brain areas in autism.

Interestingly, functional connectivity in the frontal-posterior Theory of Mind network was modulated by condition in both groups, but in different ways. Normal control participants showed an increase in functional connectivity in this network only when viewing the ToM animations, suggesting an adaptive engagement of processes involving communication among the areas only in the condition in which such processes are necessary. In contrast, the group with autism showed an increase in functional connectivity relative to the RD condition for both the GD and ToM conditions. This differential pattern of modulation resulted in the group with autism showing reliable underconnectivity relative to controls in both the RD condition and the ToM conditions, but not in the GD condition. These results suggest that individuals with autism attempt to carry out processes dependent on communication between frontal and posterior Theory of Mind areas for both types of animation involving interactions between the triangles, but that such communication is less efficient overall and less adaptive to the demands of the task.

Although the functional underconnectivity in autism is widespread in this study, affecting not only frontal-posterior connectivity but also occipital-temporal connectivity, the most intriguing facet of this is that the frontal lobe seems to be the epicenter of underconnectivity. This observation emphasizes the significant participation of the frontal lobe and its integration with other areas in a complex task, such as mental state attribution. Lower activation in frontal areas in autism has been found in several different functions, such as language comprehension (Just et al., 2004; Harris et al., 2006), verbal working memory (Koshino et al., 2005), face working memory (Koshino et al., 2008), visuospatial processing (Ring et al., 1999), and understanding emotions of others (Dapretto et al., 2006). In addition, our previous studies of executive functioning (Just et al., 2007), visual imagery (Kana et al., 2006), response inhibition (Kana et al., 2007) and discourse processing (Mason et al., 2008) have found frontal-posterior underconnectivity in autism. These findings from brain activation and functional connectivity, in conjunction with findings of structural abnormalities, such as delayed maturation of frontal lobe (Zilbovicius et al., 1995), increased frontal cortical folding (Hardan et al., 2004), anatomical shifting of major sulci in the frontal lobe (Levitt et al., 2003), and maldevelopment in minicolumns in the frontal cortex (Casanova et al., 2002; Buxhoeveden et al., 2004) suggest impairment in information transfer from and to the frontal lobe in autism.

Although the differences in frontal lobe activation between groups found for the contrast between the ToM condition and the RD condition could conceivably lead to reduced functional connectivity in the group with autism if there was simply no frontal activation in this group, note that the ROIs for which the signal intensity was extracted were not based on areas showing a group difference for this contrast, but rather were defined by the within-group contrasts between the ToM condition and the fixation baseline, a contrast for which both groups showed activation in medial frontal and bilateral inferior orbitofrontal areas. In addition, the procedure for calculating functional connectivity ignored regions within participants that did not show a sufficient volume of activation in this contrast to provide a stable estimate of the time-course of signal intensity. Given a measurable and stable signal from the involved regions of interest, our measure of functional connectivity depends on the correlation of signal change among regions during viewing of the animation, rather than on the amplitude of the signal intensity within individual regions. We also note that in previous work we have found lower functional connectivity in autism despite similar levels of activation between the groups for the regions considered (e.g., Just et al., 2007) and that such differences are found even during a resting baseline condition (Cherkassky et al., 2006). Note also that in the present study we found group differences in functional connectivity in posterior regions that did not show differences in activation between the groups. Although the present study can not distinguish whether reduced frontal activation in autism is a cause or a consequence of reduced communication between frontal and posterior Theory of Mind areas, reduced activation is not a necessary precondition for reduced functional connectivity.

Widespread white matter abnormalities have also been found in autism, including lower fractional anisotropy of the white matter adjacent to the Theory of Mind regions (medial frontal gyrus, anterior cingulate cortex, and right temporoparietal junction) (Barnea-Goraly et al., 2004), significantly larger frontal and parietal white matter volume (Carper et al., 2002), disturbances (both increases and decreases in volume in different areas) in white matter (Herbert et al., 2004), white matter volume deficits in right middle frontal and left superior frontal regions (Waiter et al., 2005), reduced thickness in orbitofrontal white matter (Chung et al., 2005), increased regional and global white matter (Hendry et al., 2006), and abnormalities in corpus callosum size (Egaas et al., 1995; Piven et al., 1997; Hardan et al., 2000). Any or all of these abnormalities could conceivably affect the functional and structural connectivity of the frontal lobe with more posterior regions in autism. These findings taken together point to abnormalities in the neural architecture in autism, which may affect several types of psychological functions, including Theory of Mind.

The critical role of the prefrontal cortex in Theory of Mind processing is supported by neuropsychological studies that show frontal lobe damage associated with impairments in Theory of Mind (Channon & Crawford, 2000; Rowe et al., 2001; Stone et al., 1998; Stuss et al., 2001), including the medial prefrontal cortex as well as the orbitofrontal cortex. Gallagher and Frith (2003) propose that while the posterior superior temporal sulcus provides the cues for mentalizing, the medial prefrontal cortex does the mental state reasoning. Therefore a communication between these two regions is vital in Theory of Mind tasks. In other words, cortical underconnectivity in autism among the components of the network for mentalizing provides a plausible neural basis of impaired Theory of Mind in autism. The underconnectivity could simply be at the level of a lack of information coordination between two areas. However, in autism, structural and abnormalities in the white matter connections to the frontal lobe have been reported by several studies, so the functional underconnectivity could well be related to anatomical connectivity properties, as we have demonstrated previously (Just et al., 2007; Kana et al., 2006; Cherkassky et al., 2006; Keller et al., 2007).

In addition to showing lower functional connectivity in autism in the Theory of Mind task, the present study also demonstrated a reduction in activation during mentalizing in three key frontal regions: the medial frontal gyrus, the orbitofrontal gyrus, and the anterior paracingulate cortex. This finding is consistent with several previous studies of Theory of Mind processing in autism (Castelli et al., 2002; Happe et al., 1996; Baron-Cohen et al., 1999). Mental state attribution presupposes self-awareness and it likely involves processing in evolutionarily recent regions of the neocortex, such as the frontal cortex (Gallup, 1982). As mentioned in the earlier part of the discussion, frontal lobe underactivation in autism is not uncommon in other higher cognitive tasks. It may be possible to distinguish the roles of frontal regions and parietal regions in the ToM network, although the differentiation is not critical to the current study, nor do the current results help to distinguish the roles. Three frontal regions have been credited with distinct roles in mental state attribution. The medial prefrontal cortex has been credited with involvement in any kind of thought that uses the self as a referent (Gusnard et al., 2001), the anterior paracingulate cortex is believed to be involved in the decoupling mechanism that distinguishes mental states from reality (Gallagher & Frith, 2003), and the orbitofrontal cortex is believed to play a crucial role in Theory of Mind (Baron-Cohen & Ring, 1994; Brothers and Ring, 1992), and may be associated with recognition of mental state terms (Baron-Cohen et al., 1994). The reduced activation in participants with autism in these regions indicates an atypical neural basis for Theory of Mind processing.

Contrary to our prediction, there was no reliable group difference in activation in the right superior temporal sulcus (STS) at the temporoparietal junction while processing Theory of Mind. This is in contrast with the findings of reduced activation in autism in this region in a previous study (Castelli et al., 2002). However, the present finding raises several interesting questions. First, in contrast to the implicit nature of the task (passively viewing the animations) in the Castelli et al. (2002) study, the present study included instructions that were designed to require participants to explicitly attempt to attribute mental states to the stimuli (watching animations and answering questions related to mental state). Normal or a higher level of activation in autism has been found in explicit tasks of processing mental states (Wang et al., 2006). It is possible that the participants with autism might make a more controlled effort to infer a mental state when they are explicitly asked to do so. Second, it is possible that the superior temporal sulcus plays a subordinate role in inferring mental states. Gallagher and Frith (2003) argued that the STS is more generally associated with social cognition but the dominant roles in Theory of Mind are played by the medial prefrontal cortex and the anterior paracingulate cortex. Third, superior temporal activation has been found during the viewing of simple interactions of moving objects (Blakemore et al., 2003; Pelphrey et al., 2005) suggesting that STS activation need not necessarily be associated with attribution of mental states. Although all regions that are found to be part of the ToM network play different roles, it may be the coordination and communication between these regions that is crucial in accomplishing Theory of Mind. As an adaptation in autism to the reduced connectivity with frontal areas, the posterior superior temporal sulcus might function with greater autonomy, assuming a higher than normal degree of control of the task performance. Thus, the present results are consistent with the idea that participants with autism may process the ToM cues in the same way as controls, but if the results of such processing are not communicated efficiently to frontal regions, then the frontal regions may not be able to normally carry out the mental state attribution. This interpretation would also predict reduced synchronization between the activation in anterior and posterior regions involved in the ToM processing in autism, which is what occurred in this study.

An interesting aspect of activation during Theory of Mind processing in autism was that the activation in the right posterior superior temporal sulcus in autism was positively correlated with the Theory of Mind score. In other words, people with autism who had higher Theory

of Mind ability also had greater activation in posterior STS. A correlation between a psychometrically measured ability and an activation level in a relevant brain area has previously been found in autism with respect to other abilities, such as the correlation between the psychometric measure of face processing ability measured by a standardized test of face perception and activation in the fusiform face area in autism (Koshino et al., in press; Schultz et al., 2005). Schultz et al. (2005) also found the intactness of social functioning in autism (measured by ADOS) to be correlated with the amount of activation in fusiform gyrus in a face processing task in people with autism, with the better-socially-functioning autism participants having more fusiform activation.

The absence of a behavioral group difference was unexpected, given previous findings of such differences and given our finding of group differences in brain activation. Note, however, that each participant's behavioral data were based on only three items in each condition. Although these few items were apparently sufficient to provide precise enough estimates to detect within-subject effects of condition, they may have been inadequate for detecting between-group differences within the ToM condition. Another possible explanation is that the explicit instructions to attend to the thoughts and feelings of the triangles and the forced choice nature of the response prevented the task from being sensitive to the well-established differences in theory of mind processing between the groups.

The present evidence for functional underconnectivity in autism in a Theory of Mind task adds to the converging evidence of functional underconnectivity in a wide range of complex information processing tasks, including language comprehension (Just et al., 2004), working memory (Koshino et al., 2005), executive functioning (Just et al., 2007), complex inhibition (Kana et al., 2007), and mental imagery (Kana et al., 2006). What is common across all these studies is a neural pattern in autism characterized by 1) frontal-posterior underconnectivity; 2) less frontal activation; 3) equivalent, or greater activation in posterior brain areas; and 4) decrease in size of a relevant portion of the corpus callosum. Disruptions in connectivity between frontal and posterior areas in autism could cause the system to adapt in such a way that the posterior regions work more independently to compensate for the impaired communication with the frontal regions. The phenomenon of frontal-posterior underconnectivity and a consequent increased posterior autonomy has been demonstrated in a computational model of a complex executive task (Just et al., in preparation). In summary, the functional connectivity and brain activation results of the present study indicate a systems-level neural impairment in people with autism. The focal point of this impairment seems to be the decreased communication and coordination between frontal regions with others, affecting Theory of Mind processing among other forms of thought.

Acknowledgments

This research was supported by the Collaborative Program of Excellence in Autism (CPEA) Grant HD35469 from the National Institute of Child Health and Human Development and the Cure Autism Now Young Investigator Award to Dr. Kana. The authors thank Dr. Fulvia Castelli (Department of Cognitive Neuroscience, Kings College, London, UK) for generously providing the animation stimuli for our study. The authors also thank Dr. Diane Williams for her assistance in collecting the ToM scores, Sarah Schipul and Stacey Becker for assistance with the data collection and data analysis, and Rachel Krishnaswami for editorial comments on an earlier version of the manuscript.

References

- Adolphs R. Cognitive neuroscience of human social behavior. *Nat Rev Neurosci.* 2003; 4:165–78. [PubMed: 12612630]
- Amodio DM, Frith CD. Meeting of minds: the medial frontal cortex and social cognition. *Nat Rev Neurosci.* 2006; 7:268–277. [PubMed: 16552413]

- Barbas H, Ghashghaei H, Dombrowski SM, Rempel-Clower NL. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory related areas in the rhesus monkey. *J Comp Neurol.* 1999; 410:343–367. [PubMed: 10404405]
- Barnea-Goraly N, Kwon H, Menon V, Eliez S, Lotspeich L, Reiss AL. White matter structure in autism: Preliminary evidence from diffusion tensor imaging. *Biol Psych.* 2004; 55:323–326.
- Baron-Cohen S, Campbell R, Karmiloff-Smith A, Grant J, Walker J. Are children with autism blind to the mentalistic significance of the eyes? *Brit J of Dev Psych.* 1995; 13:379–398.
- Baron-Cohen S, Leslie AM, Frith U. Does the autistic child have a ‘theory of mind’? *Cognition.* 1985; 21:37–46. [PubMed: 2934210]
- Baron-Cohen S, Ring H, Moriarty J, Schmitz B, Costa D, Ell P. Recognition of mental state terms. Clinical findings in children with autism and a functional neuroimaging study of normal adults. *Brit J of Dev Psych.* 1994; 165:640–649.
- Baron-Cohen, A.; Ring, H. A model of the mindreading system: neuropsychological and neurobiological perspectives. In: Mitchell, P.; Lewis, C., editors. *Origins of an understanding of mind.* Lawrence Erlbaum Associates; 1994.
- Baron-Cohen S, Ring HA, Wheelwright S, Bullmore ET, Brammer MJ, Simmons A, Williams SC. Social intelligence in the normal and autistic brain: an fMRI study. *Euro J of Neurosci.* 1999; 11:1891–1898.
- Blake R, Turner LM, Smoski MJ, Pozdol SL, Stone WL. Visual recognition of biological motion is impaired in children with autism. *Psych Sci.* 2003; 14:151–157.
- Blakemore S-J, Boyer P, Pachot-Clouard M, Meltzoff A, Segebarth C, Decety J. The detection of contingency and animacy from simple animations in the human brain. *Cereb Cor.* 2003; 13:837–844.
- Blakemore S-J, Winston J, Frith U. Social cognitive neuroscience: where are we heading? *Trends in Cog Sci.* 2004; 8:216–222.
- Brothers L. The social brain: A project for integrating primate behavior and neurophysiology in a new domain. *Concepts Neurosci.* 1990; 1:27–51.
- Brothers L, Ring B. A neuroethological framework for the representation of minds. *J of Cog Neurosci.* 1992; 4:107–118.
- Brunet E, Sarfati Y, Hardy-Bayle MC, Decety J. A PET investigation of the attribution of intentions with a nonverbal task. *NeuroImage.* 2000; 11:157–166. [PubMed: 10679187]
- Buxhoeveden DP, Semendeferi K, Schenker N, Courchesne E. Decreased cell column spacing in autism (abstract). *Social Neurosci.* 2004:582.6.
- Carmichael ST, Price JL. Connectional networks within the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol.* 1996; 371:179–207. [PubMed: 8835726]
- Carper RA, Moses P, Tigue ZD, Courchesne E. Cerebral Lobes in Autism: Early Hyperplasia and Abnormal Age Effects. *Neuroimage.* 2002; 16:1038–1051. [PubMed: 12202091]
- Casanova MF, Buxhoeveden DP, Switala AE, Roy E. Minicolumnar pathology in autism. *Neurology.* 2002; 58:428–32. [PubMed: 11839843]
- Castelli F, Frith C, Happe F, Frith U. Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain.* 2002; 125:1839–49. [PubMed: 12135974]
- Castelli F, Happé F, Frith U, Frith CD. Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage.* 2000; 12:314–325. [PubMed: 10944414]
- Channon S, Crawford S. The effects of anterior lesions on performance on a story comprehension test: left anterior impairment on a theory of mind-type task. *Neuropsychologia.* 2000; 38:1006–1017. [PubMed: 10775711]
- Cherkassky VL, Kana RK, Keller TA, Just MA. Functional connectivity in a baseline resting-state network in autism. *NeuroReport.* 2006; 17:1687–90. [PubMed: 17047454]
- Chung MK, Robbins SM, Dalton KM, Davidson RJ, Alexander AL, Evans AC. Cortical thickness analysis in autism with heat kernel smoothing. *NeuroImage.* 2005; 25(4):1256–65. [PubMed: 15850743]

- Dapretto M, Davies MS, Pfeifer JH, Scott AA, Sigman M, Bookheimer SY, Iacoboni M. Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat Neurosci*. 2006; 9:28–30. [PubMed: 16327784]
- Egaas B, Courchesne E, Saitoh O. Reduced size of corpus callosum in autism. *Arch of Neurol*. 1995; 52:794–801. [PubMed: 7639631]
- Fletcher PC, Happé F, Frith U, Baker SC, Dolan RJ, Frackowiak RSJ, Frith CD. Other minds in the brain: a functional imaging study of ‘theory of mind’ in story comprehension. *Cognition*. 1995; 57:109–128. [PubMed: 8556839]
- Friston KJ, Holmes AP, Worsley KJ, Poline J-B, Frith CD, Frackowiak RSJ. Statistical parametric maps in functional imaging: A general linear approach. *Hum Br Map*. 1995; 2:189–210.
- Frith U. Mindblindness and the brain in autism. *Neuron*. 2001; 32:969–979. [PubMed: 11754830]
- Gallagher HL, Frith CD. Functional imaging of ‘theory of mind’. *Trends in Cog Sci*. 2003; 7:77–83.
- Gallagher HL, Jack AI, Roepstorff A, Frith CD. (2002) Imaging the intentional stance in a competitive game. *Neuroimage*. 2002; 16:814–21. [PubMed: 12169265]
- Gallagher HL, Happe F, Brunswick N, Fletcher PC, Frith U, Frith CD. Reading the mind in cartoons and stories: an fMRI study of ‘theory of mind’ in verbal and nonverbal tasks. *Neuropsychologia*. 2000; 38:11–21. [PubMed: 10617288]
- Gallup GG Jr. Self-awareness and the emergence of mind in primates. *Am J of Primatology*. 1982; 2:237–248.
- Gusnard DA, Akbudak E, Shulman GL, Raichle ME. Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc Natl Acad Sci USA*. 2001; 98:4259–4264. [PubMed: 11259662]
- Happe FG. An advanced test of theory of mind: understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *J of Autism and Dev Disord*. 1994; 24:129–154. [PubMed: 8040158]
- Happe FG, Ehlers S, Fletcher P, Frith U, Johansson M, Gillberg C, Dolan R, Frackowiak R, Frith C. 'Theory of mind' in the brain. Evidence from a PET scan study of Asperger syndrome. *Neuroreport*. 1996; 8:197–201. [PubMed: 9051780]
- Hardan AY, Minshew NJ, Keshavan MS. Corpus Callosum Size in Autism. *Neurology*. 2000; 55(7): 1033–6. [PubMed: 11061265]
- Hardan AY, Jou RJ, Keshavan MS, Varma R, Minshew NJ. Increased frontal cortical folding in autism: a preliminary MRI study. *Psychiatry Res*. 2004; 131:263–68. [PubMed: 15465295]
- Harris GH, Chabris CF, Clark J, Urban T, Aharon I, Steele S, McGrath L, Condouris K, Tager-Flusberg H. Brain activation during semantic processing in autism spectrum disorders via functional magnetic resonance imaging. *Br and Cog*. 2006; 61:54–68.
- Herbert MR, Ziegler DA, Deutsch CK, O'Brien LM, Lange N, Bakardjiev A, Hodgson J, Adrien KT, Steele S, Makris N, Kennedy D, Harris GJ, Caviness VS Jr. Dissociations of cerebral cortex, subcortical and cerebral white matter volumes in autistic boys. *Brain*. 2003; 126:1182–1192. [PubMed: 12690057]
- Herbert MR, Ziegler DA, Makris N, Filipek PA, Kemper TL, Normandin JJ, et al. Localization of white matter volume increase in autism and developmental language disorder. *Ann Neurol*. 2004; 55:530–540. [PubMed: 15048892]
- Hendry J, DeVito T, Gelman N, Densmore M, Rajakumar N, Pavlosky W, Williamson PC, Thompson PM, Drost DJ, Nicolson R. White matter abnormalities in autism detected through transverse relaxation time imaging. *Neuroimage*. 2006; 29:1049–57. [PubMed: 16214373]
- Jackson PL, Brunet E, Meltzoff AN, Decety J. Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia*. 2006; 44:752–761. [PubMed: 16140345]
- Just MA, Cherkassky VL, Keller TA, Minshew NJ. Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain*. 2004; 127:1811–21. [PubMed: 15215213]
- Just MA, Cherkassky VL, Keller TA, Kana RK, Minshew NJ. Functional and anatomical cortical underconnectivity in autism: evidence from an fMRI study of an executive function task and corpus callosum morphometry. *Cereb Cor*. 2007; 17:951–961.

- Kana RK, Keller TA, Minshew NJ, Just MA. Inhibitory control in high-functioning autism: decreased activation and underconnectivity in inhibition networks. *Biol Psych*. 2007; 62:198–206.
- Kana RK, Keller TA, Cherkassky VL, Minshew NJ, Just MA. Sentence comprehension in autism: thinking in pictures with decreased functional connectivity. *Brain*. 2006; 129:2484–2493. [PubMed: 16835247]
- Keller TA, Kana RK, Just MA. A developmental study of the structural integrity of white matter in autism. *Neuroreport*. 2007; 18:23–27. [PubMed: 17259855]
- Koshino H, Carpenter PA, Minshew NJ, Cherkassky VL, Keller TA, Just MA. Functional connectivity in an fMRI working memory task in high-functioning autism. *NeuroImage*. 2005; 24:810–821. [PubMed: 15652316]
- Leekam S, Perner J. Does the autistic child have a metarepresentational deficit? *Cognition*. 1991; 40:203–218. [PubMed: 1786675]
- Levitt JG, Blanton RE, Smalley S, Thompson PM, Guthrie G, McCracken JT, Sadoun T, Heinichen L, Toga AW. Cortical sulcal maps in autism. *Cereb Cor*. 2003; 13:728–735.
- Lord C, Risi S, Lambrecht L, Cook EH Jr, Leventhal BL, DiLavore PC, Pickles A, Rutter M. The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *J of Autism and Dev Disord*. 2000; 30:205–23. [PubMed: 11055457]
- Lord C, Rutter M, LeCouteur A. Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J of Autism and Dev Disord*. 1994; 24:659–85. [PubMed: 7814313]
- Minshew NJ, Goldstein G, Siegel DJ. Neuropsychological Functioning in Autism: Profile of a Complex Information Processing Disorder. *J of the Intl Neuropsych Soc*. 1997; 3:303–16.
- Pelphrey KA, Morris JP, Michelich CR, Allison T, McCarthy G. Functional anatomy of biological motion perception in posterior temporal cortex: an fMRI study of eye, mouth and hand movements. *Cereb Cor*. 2005; 15:1866–1876.
- Perner J, Frith U, Leslie AM, Leekam S. Exploration of the autistic child's theory of mind: knowledge, belief, and communication. *Ch Dev*. 1989; 60:689–700.
- Piven J, Bailey J, Ranson BJ, Arndt S. An MRI study of the corpus callosum in autism. *Am J Psych*. 1997; 154:1051–56.
- Reed T, Peterson C. A comparative study of autistic subjects' performance at two levels of visual and cognitive perspective taking. *J of Autism and Dev Disord*. 1990; 20:555–568. [PubMed: 2279973]
- Ring HA, Baron-Cohen S, Wheelwright S, Williams SCR, Brammer M, Andrew C, Bullmore ET. Cerebral Correlates of Preserved Cognitive Skills in Autism. *Brain*. 1999; 122:1305–1315. [PubMed: 10388796]
- Rowe AD, Bullock PR, Polkey CE, Morris RG. 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*. 2001; 124:600–616. [PubMed: 11222459]
- Ruby P, Decety J. Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat Neurosci*. 2001; 4:546–50. [PubMed: 11319565]
- Sabbagh MA. Understanding orbitofrontal contributions to theory-of-mind reasoning: implications for autism. *Br and Cog*. 2004; 55:209–219.
- Saxe R, Kanwisher N. People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". *Neuroimage*. 2003; 19(4):1835–42. [PubMed: 12948738]
- Schultz, RT.; Win, L.; Jackowski, A.; Klin, A.; Staib, L.; Papademetris, X.; Babitz, T.; Carter, E.; Klaiman, C.; Feiler, A.; Volkmar, F. Brain morphology in autism spectrum disorders: an MRI study. Paper presented at annual Intl. Mtg for Autism Res; Boston, MA. May 2005;
- Stone VE, Baron-Cohen S, Knight RT. Frontal lobe contributions to theory of mind. *J of Cog Neurosci*. 1998; 10:640–656.
- Stuss, DT. Self, awareness and the frontal lobes: a neuropsychological perspective. In: Strauss, J.; Goethals, GR., editors. *The Self: interdisciplinary approaches*. New York: Springer-Verlag; 1991. p. 255-78.
- Stuss DT, Gallup GG, Alexander MP. The frontal lobes are necessary for theory of mind. *Brain*. 2001; 124:279–286. [PubMed: 11157555]

- Swettenham J. Can children with autism be taught to understand false belief using computers? *J of Ch Psych and Psychiatry*. 1996; 37:157–165.
- Swettenham J, Baron-Cohen S, Gomez JC, Walsh S. What's inside a person's head? Conceiving of the mind as a camera helps children with autism develop an alternative theory of mind. *Cog Neuropsychiatry*. 1996; 1:73–88.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*. 2002; 15:273–89. [PubMed: 11771995]
- Villalobos ME, Mizuno A, Dahl BC, Kemmotsu N, Muller RA. Reduced functional connectivity between V1 and inferior frontal cortex associated with visuomotor performance in autism. *Neuroimage*. 2005; 25:916–25. [PubMed: 15808991]
- Vogeley K, Bussfeld P, Newen A, Herrmann S, Happe F, Falkai P, Maier W, Shah NJ, Fink GR, Zilles K. Mind reading: neural mechanisms of theory of mind and self-perspective. *NeuroImage*. 2001; 14:170–181. [PubMed: 11525326]
- Waiter GD, Williams JHG, Murray AD, Gilchrist A, Perrett DI, Whiten A. (2005) Structural white matter deficits in high-functioning individuals with autistic spectrum disorder: a voxel-based investigation. *Neuroimage*. 2005; 24:455–61. [PubMed: 15627587]
- Wang AT, Lee SS, Sigman M, Dapretto M. Neural basis of irony comprehension in children with autism: the role of prosody and context. *Brain*. 2006; 129:932–943. [PubMed: 16481375]
- Winston JS, Strange BA, O'Doherty J, Dolan RJ. Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat Neurosci*. 2002; 5:277–83. [PubMed: 11850635]
- Witelson SF. Hand and sex differences in the isthmus and genu of the human corpus callosum: a postmortem morphological study. *Brain*. 1989; 112:799–835. [PubMed: 2731030]
- Zilbovicius M, Garreau B, Samson Y, Remy P, Barthelemy C, Syrota A, Lelord G. Delayed Maturation of the Frontal Cortex in Childhood Autism. *Am J of Psychiatry*. 1995; 152:248–52. [PubMed: 7840359]

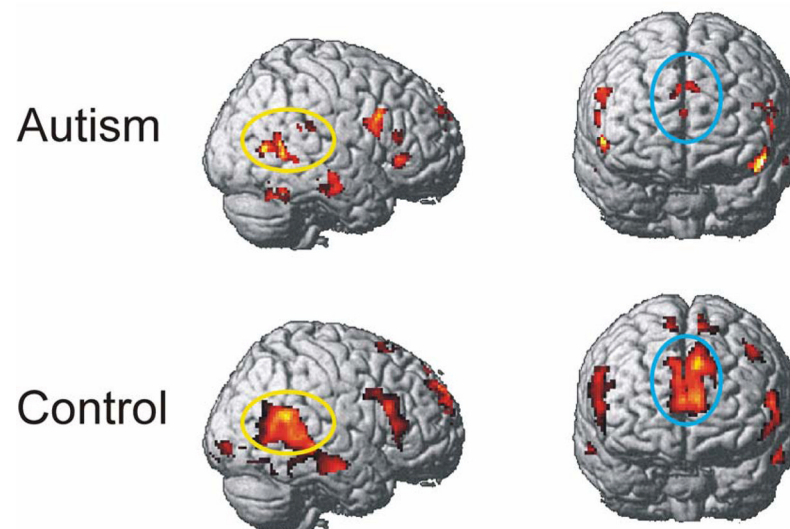


Figure 1. Activation in right superior temporal (yellow ovals) and superior medial frontal (blue ovals) areas in autism and control groups for the contrast of Theory of Mind processing with Random animation processing. The t-maps are thresholded at $p < .005$ and an extent threshold of 6 voxels.

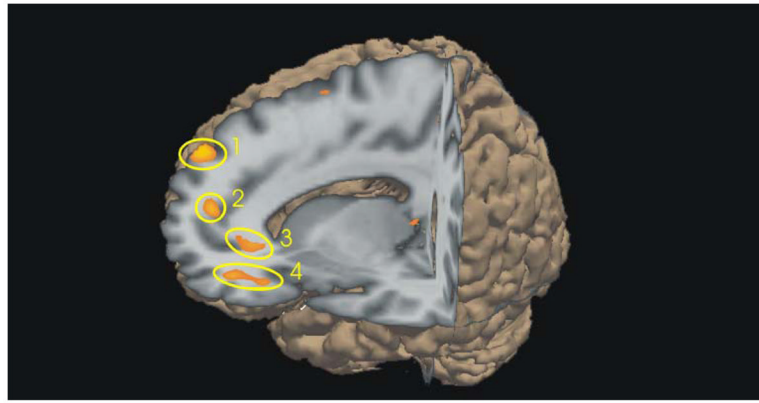


Figure 2. Participants with autism showed reduced levels of activation relative to control participants ($p < .005$, uncorrected) in four regions in the frontal Theory of Mind network: 1) left superior medial frontal gyrus, 2) left anterior paracingulate cortex, 3) bilateral anterior cingulate cortex, and 4) left inferior orbitofrontal cortex. The graphic was created using Slicer software: <http://slicer.org>.

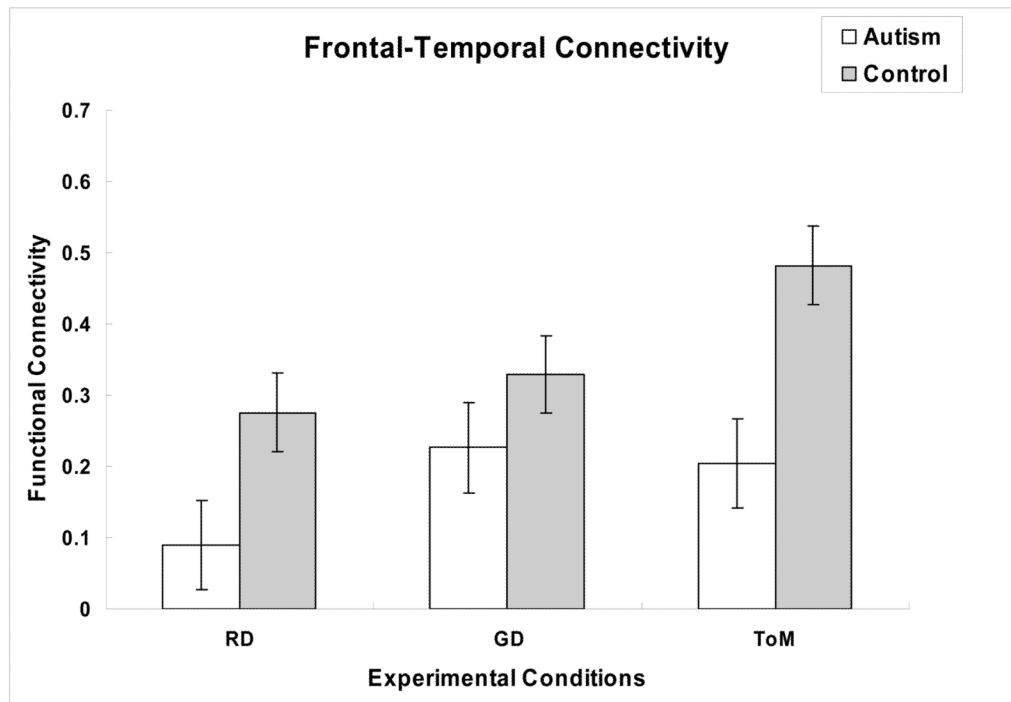


Figure 3. Mean functional connectivity (z') in the frontal-posterior Theory of Mind network for each group and each condition. Participants with autism showed significantly reduced functional connectivity relative to control participants in the Theory of Mind condition and the Random condition, but not in the Goal-directed condition.

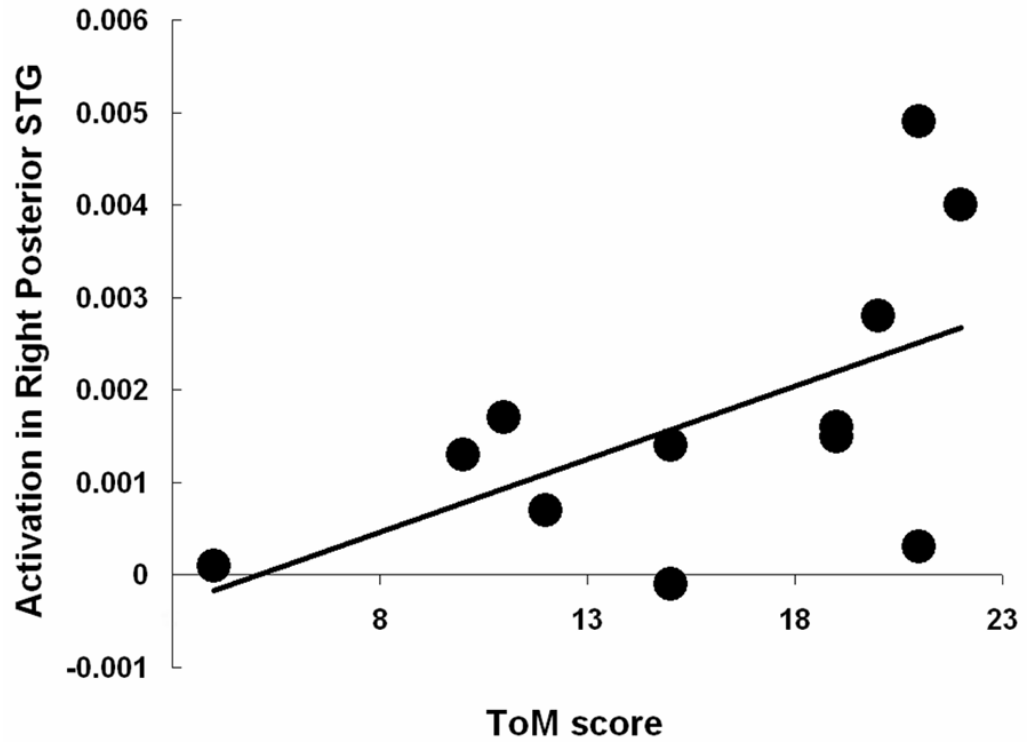
Autism Group ($r = 0.57$)

Figure 4. Participants with autism showed a significant positive correlation between Theory of Mind score and the activation in the right posterior superior temporal gyrus.

Table 1

Age, IQ, handedness, and gender of participants

		Autism	Control
Age (years)	Range; Mean \pm SD	15.8–35.7; 24.6 \pm 6.9	16.6–31.4; 24.4 \pm 3.7
Verbal IQ	Range; Mean \pm SD	81–135; 104.1 \pm 14.4	89–126; 110.6 \pm 9.9
Performance IQ	Range; Mean \pm SD	75–129; 102.8 \pm 13.6	100–127; 115.2 \pm 7.4
Full Scale IQ	Range; Mean \pm SD	83–136; 104.3 \pm 14.4	99–129; 114.3 \pm 8.6
Handedness	Right : left	11 : 1	11 : 1
Gender	Male : female	10 : 2	12 : 0

Note. IQ: Intelligence Quotient; SD: Standard Deviation

Table 2

Brain activation within autism and control groups when theory of mind processing was contrasted with processing of random animations.

Location of peak activation	Brodmann's Area	Cluster Size	t(11)	MNI Coordinates		
				X	y	z
<i>Participants with autism</i>						
R inferior frontal gyrus	47	25	7.63	52	30	0
L inferior frontal-orbital	47	51	7.46	-48	38	-14
R fusiform gyrus	39	49	6.17	42	-54	-24
R inferior frontal gyrus	9	46	5.42	46	8	24
R middle temporal gyrus	22	28	4.97	58	-52	6
<i>Typical control participants</i>						
L superior medial frontal gyrus	9	1037	12.37	-10	56	42
L angular gyrus	39	1009	11.16	-50	-72	28
L inferior frontal gyrus	45	228	8.18	-54	20	8
R middle temporal gyrus	13	1006	7.70	44	-48	18
R precuneus	7	407	7.08	0	-62	32
L middle frontal gyrus	8	43	7.01	-42	18	48
R middle temporal gyrus	21	172	6.91	60	-10	-18
R fusiform gyrus	37	25	6.22	44	-46	-22
L middle temporal gyrus	21	113	6.17	-54	-12	-24
L fusiform gyrus	37	125	6.01	-42	-56	-14
L middle temporal gyrus	20	33	5.35	-54	-34	-10
R inferior frontal gyrus	45	87	5.30	60	24	22

Note: L, left; R, right. Entries in bold are significant at $p < .05$ corrected for multiple comparisons on the basis of the cluster-size. Entries in light type are significant at $p < .001$, uncorrected. Region labels, t-values, and MNI coordinates are for the peak activated voxel in each cluster only.

Table 3

Functional connectivity differences in interlobe and intralobe networks between autism and control groups for the Theory of Mind condition.

Interlobe Networks	Autism	Control	Group Difference	F value	p value
Frontal-Parietal	0.21	0.51	0.30	8.60	0.008
Frontal-Occipital	0.23	0.56	0.33	11.89	0.003
Frontal-Temporal	0.32	0.57	0.25	7.75	0.012
Parietal-Temporal	0.45	0.68	0.23	5.68	0.026
Parietal-Occipital	0.48	0.76	0.28	6.95	0.015
Occipital-Temporal	0.70	0.91	0.21	4.35	0.049
Intralobe Networks	Autism	Control	Group Difference	F value	p value
Frontal-Frontal	0.28	0.49	0.21	3.74	0.068
Temporal-Temporal	0.58	0.83	0.25	3.27	0.086
Parietal-Parietal	0.53	0.82	0.29	5.65	0.027
Occipital-Occipital	0.67	0.95	0.28	5.81	0.025