January 2008

Theory of mind disruption and recruitment of the right hemisphere during narrative comprehension in autism

Robert A. Mason  
*Carnegie Mellon University*

Diane L. Williams  
*Duquesne University*

Rajesh K. Kana  
*University of Alabama - Birmingham*

Nancy J. Minshew  
*University of Pittsburgh School of Medicine*

Marcel Adam Just  
*Carnegie Mellon University, just@cmu.edu*

Follow this and additional works at: [http://repository.cmu.edu/psychology](http://repository.cmu.edu/psychology)

Part of the Artificial Intelligence and Robotics Commons, Cognition and Perception Commons, Cognitive Neuroscience Commons, Cognitive Psychology Commons, Computational Neuroscience Commons, Developmental Neuroscience Commons, Discourse and Text Linguistics Commons, First and Second Language Acquisition Commons, and the Semantics and Pragmatics Commons

**Published In**  
*Neuropsychologia*, 269-280.
Theory of Mind disruption and recruitment of the right hemisphere during narrative comprehension in autism

Robert A. Mason a,∗, Diane L. Williams a,b, Rajesh K. Kana a, Nancy Minshew c, Marcel Adam Just a

a Center for Cognitive Brain Imaging, Carnegie Mellon University, United States
b Department of Speech-Language Pathology, Duquesne University, United States
c Departments of Psychiatry and Neurology, University of Pittsburgh School of Medicine, United States

Received 4 January 2007; received in revised form 26 June 2007; accepted 18 July 2007
Available online 1 August 2007

Abstract

The intersection of Theory of Mind (ToM) processing and complex narrative comprehension in high functioning autism was examined by comparing cortical activation during the reading of passages that required inferences based on either intentions, emotional states, or physical causality. Right hemisphere activation was substantially greater for all sentences in the autism group than in a matched control group suggesting decreased LH capacity in autism resulting in a spillover of processing to RH homologs. Moreover, the ToM network was disrupted. The autism group showed similar activation for all inference types in the right temporo-parietal component of the ToM network whereas the control participants selectively activated this network only when appropriate. The autism group had lower functional connectivity within the ToM network and also between the ToM and a left hemisphere language network. Furthermore, the within-network functional connectivity in autism was correlated with the size of the anterior portion of the corpus callosum.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Language; Inferences; fMRI; Functional connectivity; Corpus callosum; Cortical networks

Autism involves primary impairments in language comprehension and pragmatics, or the ability to use language to communicate effectively in social contexts (Lord & Paul, 1997; Tager-Flusberg, 1981, 1996; Wilkinson, 1998). These pragmatic language impairments are thought to be related to deficits in Theory of Mind, or the capacity to make inferences about what others think, feel and know (Baron-Cohen, Leslie, & Frith, 1985; Baron-Cohen, 1988; Happé, 1993; Tager-Flusberg, 1993, 1997). Several studies have found atypical brain activation in autism during the processing of language (Harris et al., 2006; Just, Cherkassky, Keller, & Minshew, 2004; Kana, Keller, Cherkassky, Minshew, & Just, 2006) and during mental state attribution or Theory of Mind tasks (Castelli, Frith, Happé, & Frith, 2002; Schultz et al., 2003). Furthermore, a recent study on the comprehension of irony found significantly higher activation in Theory of Mind regions in autism, indicating the difficulty participants with autism face in tasks that involve language and Theory of Mind (Wang, Lee, Sigman, & Dapretto, 2006).

The interrelationship of the processes of language comprehension and Theory of Mind in individuals with autism may profitably be examined within the context of discourse processing, in which understanding the intentionality of the protagonist plays a central role (Gernsbacher, Hallada, & Robertson, 1998). Neuroimaging research with typical individuals suggests that whereas there are many lower level processes involved in discourse comprehension (such as language processing at the word and sentence levels), discourse critically involves processing at higher levels (Ferstl, 2006; Ferstl, Neumann, Bogler, & von Cramon, 2007). We have proposed a model of approximately five Parallel Networks of Discourse (Mason & Just, 2006) that process figurative and meta-sentence level information during discourse comprehension. These networks include: a coarse semantic processing network (right middle and superior temporal), a coherence monitoring network (bilateral dorso-
lateral prefrontal), a text integration network (left inferior frontal–left anterior temporal), a spatial imagery network (left dominant, bilateral intraparietal sulcus), and most relevant for the current study, a network for interpreting a protagonist’s or agent’s perspective (bilateral medial frontal/posterior right temporal/parietal). This last network effectively applies Theory of Mind processes to the comprehension of a narrative. According to this view discourse processing, at the neural level, requires coordination and communication among several brain regions involved in language and Theory of Mind.

Theory of Mind (understanding the thoughts of another person) and many “social interpretation processes” may be used in discourse processing to understand protagonists’ actions (Castelli et al., 2002; Gernsbacher et al., 1998). Consider for example this set of sentences:

Brad had no money but he just had to have the beautiful ruby ring for his wife. Seeing no salespeople around, he quietly made his way closer to the ring on the counter. He was seen running out the door.

The sentence invites the inference that Brad stole the ring. The reader can make this inference based on the information about the manner in which the protagonist approached a valuable object in a retail environment. Theory of Mind or an understanding of others’ minds is clearly necessary for interpreting the intentions, goals, and actions of characters within a narrative.

The primary goal of the current study was to determine how the cortical networks that underlie discourse processing operate in individuals with autism compared to matched control individuals. In particular, considering the difficulty individuals with autism have in performing tasks that require the Theory of Mind processing, the protagonist-monitoring network might be expected to operate suboptimally. An impairment in Theory of Mind processing in individuals with autism (e.g., Happé, 1994) should affect the comprehension of texts that invite interpretation of the intention, goals, and actions of the characters and should be detectable with functional magnetic resonance imaging. In addition to activation differences, one would expect the autism group to have lower measures of functional connectivity than the control group both within the Theory of Mind network and between this network and other networks during discourse processing. This prediction arises because functional underconnectivity between the frontal lobe and other regions has been found in autism in language comprehension (Just et al., 2004; Kana et al., 2006) and Theory of Mind (Castelli et al., 2002) tasks. In addition, the functional connectivity between the frontal and parietal has repeatedly been found to be abnormally low (Cherkassky, Kana, Keller, & Just, 2006; Just, Cherkassky, Keller, Kana, & Minshew, 2007). This particular interregional underconnectivity is relevant here because the Theory of Mind network includes frontal and parietal areas.

In addition to examining Theory of Mind processing in discourse, the generality and specificity of the impairment associated with autism in other aspects of discourse processing were explored. In other words, if there is an impairment in discourse processing in autism, is it limited to inferences that are based on human intentionality (based on ToM), or does it extend to inferences based on other types of information? Prior behavioral research has indicated that individuals with high functioning autism performed as well as controls on discourse that involves inferences about physical states, but performed statistically reliably worse than the control groups on inferences about mental states (Happé, 1994; Jolliffe & Baron-Cohen, 1999). In addition, the relationship between understanding of emotional states and Theory of Mind has been of interest in autism (Baron-Cohen et al., 1999). Therefore, the current study investigated three types of inferences: intentional, physical, and emotional. An example of a passage which invites an inference based on physical causality would be:

Jane knew Pete loved pot roast so she invited him over for dinner. While the pot roast was in the oven, the telephone rang. Jane had to open up all her windows to let the smoke out.

Here the relation between roasting and smoke is based on physical causality rather than human intentionality. The third type of inference, based on understanding of emotions, entails an inference about the emotional state of the protagonist such as:

Stacy was very sad because her grandmother died yesterday. After Stacy told Jen how sad she was, Jen kept thinking about Stacy. Jen baked Stacy cookies and went to visit her.

In this example, the reader has to infer Jen’s emotional state of sympathy as being the source of the action. Based on prior behavioral research, the likelihood of impairment of discourse processing in the participants with autism was predicted to be highest for intentional (ToM-based) inferences, intermediate for emotional inferences, and lowest for physical inferences. An example of each type of passage is provided in Table 1.

*Discourse processing also allows the examination of the involvement of the right hemisphere during language processing in autism. A number of researchers have noted the similarities between the pragmatic and the discourse processing problems of individuals with autism and those of individuals with right hemispheric brain damage. This similarity leads to the hypothesis that individuals with autism may not engage the right hemisphere to the same degree as control participants during language processing (Ellis, Ellis, Fraser, & Deb, 1994; Ozonoff & Miller, 1996; Sabbagh, 1999; Shields, Varley, Broks, & Simpson, 1996). However, a recent study of text comprehension showed that people with high-functioning autism had *higher* activation in right temporal regions while processing scenarios involving irony (Wang et al., 2006). It may be that making inferences in discourse processing is more difficult in autism for a variety of reasons, all of which cumulate to produce a larger processing load than the left-hemisphere-dominant language networks can handle. Several behavioral studies have previously shown that individuals with autism have difficulty using contextual cues to make appropriate inferences, establishing that this is a difficult task for this group (Dennis, Purvis, Barnes, Wilkinson, & Winner, 2001; Minshew, Goldstein, & Siegel, 1995; Ozonoff & Miller, 1996). As a result of the extra difficulty, excessive processing demands may consume the resources of the LH language areas and the processing may spill over from the primarily left-hemisphere processing
centers to their less specialized right hemisphere homologs (Just, Carpenter, Keller, Eddy, & Thulborn, 1996; Just, Carpenter, & Just, 1999). In the case of individuals with autism, there could be a more general involvement of the right hemisphere while reading connected discourse as a result of spillover of processing from the left hemisphere areas. When people with autism face more difficulties than controls in comprehending discourse, they may meet that difficulty with RH activation.

It is important to consider the structural properties of the brain when describing the functional connectivity and to relate the structural properties to the functional properties revealed by brain activation measures. Previous research has found a correlation between the size of regions of the corpus callosum and the functional connectivity between the cortical regions connected through these regions (Just et al., 2007; Kana et al., 2006). To examine this structure-function correlation in discourse processing, the size of the various segments of the corpus callosum of each participant in the current discourse processing study was measured. Additionally, as in previous studies, the correlation between the size of various corpus callosum segments and the functional connectivity between the critical networks that they connect was examined. The prediction was that in the participants with autism, there would be a positive correlation, that is, the lower the functional connectivity, the smaller the segment. The prediction is based on the assumption that the corpus callosum size in autism is an index of the integrity of its function, and that a smaller corpus callosum puts more constraint on the functional connectivity.

1. Experimental procedures
1.1. Participants

The participants were 18 high-functioning individuals with autism (17 males and 1 female) and 18 healthy normal control participants (16 males and 2 females). They were matched for age (autism group mean: 26.5 and control group mean: 27.4, t(34) = 0.24, ns), Full Scale, Verbal IQ, and Performance IQ as determined by the Wechsler Adult Intelligence Scale Revised (WAIS-R, 80 or above), and gender. The average Full Scale IQ was 101.9 for the autism group and 105.5 for the control group, t(34) = 0.91, ns, the average Verbal IQ was 99.2 for the autism group and 103.3 for the control group, t(34) = 1.07, ns, and the average Performance IQ was 104.3 for the autism group and 106.3 for the control group, t(34) = 0.43, ns.

The diagnosis of autism was established using two structured research diagnostic instruments, the ADI-R (Autism Diagnostic Interview-Revised, Lord, Rutter, & LeCouteur, 1994) and the ADOS-G (Autism Diagnostic Observation Schedule-General, Lord et al., 2000), supplemented with expert clinical opinion according to accepted criteria of high-functioning autism (Minshew, 1996). Potential participants with autism were excluded if they had an identifiable cause for their autism such as fragile-X syndrome, tuberous sclerosis, or fetal cytomegalovirus infection. Potential control and autism participants were also excluded if there was evidence of birth asphyxia, head injury, or a seizure disorder. Exclusionary criteria were based on neurologic history and examination, and chromosomal analysis, or metabolic testing if indicated.

The control participants were community volunteers recruited to match the autism participants on age, Full Scale IQ, gender, race, and socioeconomic status of family of origin, as measured by the Hollingshead method (Hollingshead, 1957). Potential control participants were screened by questionnaire, telephone, face-to-face interview, and observation during screening psychometric tests. 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, substance abuse, 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 a family history (in parents, siblings, and offspring) of autism, developmental cognitive disorders, affective disorders, anxiety disorders, schizophrenia, obsessive compulsive disorder, substance abuse, or other neurologic or psychiatric disorder thought to have a genetic component. Handedness was determined with the Lateral Dominance Examination from the Halstead-Reitan Neuropsychological Test Battery (Reitan, 1985), revealing that three participants with autism were left-handed. The brain activation data from these left-handers were clearly similar to their respective groups, and therefore, the data were not separated by handedness.

1.2. Materials and procedure

The task of the participants was to read 30 three-sentence stories (10 passages in each of the inference versions of physical, intentional and emotional) and to respond to a simple yes/no comprehension question. The first two sentences provided a context for the passage. The second sentence of the context made it possible to generate a predictive inference that an event would occur later in the story. The third sentence constituted a potential coherence gap (this can alternatively be viewed as the cohesiveness of the text), which could be bridged through a causal inference. In the physical passages the causal inference involved a direct consequence; in the intentional passages the causal inference was guided by a character’s goal; in the emotional passages the causal inference was related to the character’s emotion. Examples of all three types are given in Table 1. The
The data were collected using a Siemens Allegra 3.0T scanner at the Brain Imaging Research Center (BIRC) of Carnegie Mellon University and the University of Pittsburgh. The study was performed with a gradient echo, EPI sequence (Tzourio-Mazoyer et al., 2002). A sphere was defined for each cluster (with a radius ranging from 5 to 10 mm) that best captured the cluster of activation in the sample intentional passage provided in the Murray, Klin, and Myers materials. Each experimental trial began with the two-sentence context. The first two context sentences appeared for 14 s (the onset was time locked to the acquisition of the superior most slice in prescription). A rest ‘X’ appeared on the screen for 4 s. The critical causal inference/control sentence region then replaced the ‘X’ and remained on the screen for 7 s fixation followed by a second rest ‘X’ for 4 s. Two passages in each condition ended with a question with two answer choices below it. The participant was given 6 s to answer the question. An ‘X’ then appeared for a 3 s rest period. The purpose of the question was to confirm that the participant comprehended the passages. Half of the answers to the questions were “yes” and half were “no”. An example of this entire sequence for a passage with a question is shown in Fig. 1.

An “X” was presented for three 30-s long fixation periods, which were interspersed among the test items, one at the beginning, one after 16 trials, and then one at the end. The participants were instructed to “please just relax, clear your mind and wait for the next story to appear” at the end of each story. Prior to entering the scanner, the participants completed a short practice set of six items, two of each kind, to familiarize them with the task.

1.3. fMRI procedures

The data were collected using a Siemens Allegra 3.0T scanner at the Brain Imaging Research Center (BIRC) of Carnegie Mellon University and the University of Pittsburgh. The study was performed with a gradient echo, EPI sequence with TR = 1000 ms, TE = 30 ms and a 60° flip angle. Sixteen oblique-axial slices were acquired; each slice was 5-mm thick with a gap of 1-mm between slices. The acquisition matrix was 64 × 64 with 3.125-mm × 3.125 × 5-mm voxels. A 160-slice axial 3D MPRAGE volume scan with TR = 2000 ms, TE = 3.34 ms, flip angle = 7°, FOV = 256 cm and a 256 × 256 matrix size, was acquired for each participant to be used in segmenting the corpus callosum into anatomically predefined regions.

1.4. fMRI analyses—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, motion-corrected, normalized to the Montreal Neurological Institute (MNI) template, resampled to 2 mm × 2 mm × 2 mm voxels, and smoothed with an 8-mm Gaussian kernel to decrease spatial noise. Statistical analysis was performed on individual and group data by using the general linear model and Gaussian random field theory as implemented in SPM99 (Friston et al., 1995). A separate regressor was created for the context and the inference sentences (as well as the questions) by convolving a boxcar function with the standard hemodynamic response function as specified in SPM. Group analyses were performed using a random-effects model. Statistical maps were superimposed on normalized T1-weighted images. An uncorrected height threshold of \( t = 3.36 \) (\( P = 0.001 \)) and an extent threshold of 6–8 mm\(^3\) voxels was used. Contrasts of the control group and the autism group were conducted on the group activation map for all three of the inference sentence contrasts versus fixation (Physical, Intentional, and Emotional), the context sentences versus fixation, as well as an analysis which collapsed across all three passage types for the context sentences.

1.5. fMRI analyses—functional connectivity

The functional connectivity was computed (separately for each participant) as a correlation between the average time course of signal intensity of all the activated voxels in each member of a pair of ROIs. Eleven functional ROIs were defined to encompass the main clusters of activation in the group activation map for each group in all three of the inference sentence contrasts versus fixation (Physical, Intentional, and Emotional). Labels for these 11 ROIs [the left medial frontal gyrus (LMedFG), the left precentral gyrus (Lprecen), the left inferior frontal gyrus (IFG) plus four bilateral ROIs, namely middle temporal gyrus (MTG), the anterior middle temporal gyrus (MTGA) the inferior occipital gyrus (IOG), and the tempo-parietal junction (TPJ)] 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). A sphere was defined for each cluster (with a radius ranging from 5 to 10 mm) that best captured the cluster of activation in the map for each group. The ROIs used in the analysis were the union of the three spheres defined for the two groups. The activation time course extracted for each participant over the activated voxels within the ROI originated from the normalized and smoothed images, which were high-pass filtered and had the linear trend removed. Participants who did not have activation in a given functional ROI were excluded from further analysis involving that ROI. The number
of participants dropped varied across ROIs; the average number of dropped participants for the control group was 1.6 and for the autism group was 3.4. The number of participants is reported below for each of the reported contrasts. The correlation was computed on the images belonging only to the inference sentence in the 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 all three conditions as well as the combined condition (collapsing across the three passage types for the inference sentence) using the 11 functional ROIs described above. There were 54 pairs of ROIs that resulted from the pairing of the 11 ROIs considered. Two sample t-tests were computed for each of the 216 ROI pairs (54 pairs for each of the three inference types and the combination of inference types). The test was one-tailed with a P < 0.05, based on the expectations that individuals with autism show lower functional connectivity than controls (Castelli et al., 2002; Just et al., 2004, 2007; Kana et al., 2006; Koshino et al., 2005). In a secondary network level analysis, several functional networks were created by combining ROIs that were thought to be working together as part of a network. Based on a proposal in Mason and Just (2006), five non-independent groupings of ROIs relevant to this task were created: (1) Lang – a left hemisphere language network containing LIFG, and LMT; (2) RLang – a right hemisphere language network consisting of RMT (a single ROI); (3) ToM – a Theory of Mind network containing LMedFG and RSTPJ; (4) TxtInt – a left hemisphere language network containing LIFG, and LMT; (5) Occi – a low-level vision network containing bilateral OCCL. Some variations of these networks were examined, including a set in which the networks were mutually exclusive. The results were similar across variations.

1.6. 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 subregions of the corpus callosum include the rostrum, genu, rostral body, anterior midbody, posterior midbody, isthmus, and splenium. In addition, there were two linear measurements: the distance between the anterior-most and posterior-most point, which is considered the length of the corpus callosum, and the callosal width at the midpoint. The corpus callosum size was normalized by dividing by the total gray and white matter volume (total cerebral volume) for each participant. The gray matter, white matter and cerebrospinal fluid volumes were measured for each subject by segmenting the T1-weighted structural brain image into three masks. The segmentation was performed by SPM2 routine. The outer contour of the corpus callosum was manually traced (with an inter-rater reliability of 0.87), and then interior segmentation, area, and length computations were performed by image processing software.

Corpus callosum measurements were compared for 17 of the 18 participants with autism and 17 of the 18 controls; one member of each group requested to be removed from the scanner after functional scans but before the structural scans (all reported r-values are significant with 32 d.f. and an alpha level of <0.05). The corpus callosum measurements (in mm²) were then transformed into a ratio with respect to an individual’s overall brain size (in mm³).

2. Results

2.1. Overview

Participants with high-functioning autism showed more right hemisphere activation than did matched controls during discourse processing. This greater activation in the right hemisphere suggests that individuals with autism found it more difficult to comprehend the passages than did the control group. Whereas the control participants recruited a region involved in Theory of Mind processing (the right temporo-parietal junction) only when it was appropriate, the participants with autism recruited this area even while processing physical inferences. Hence, there is a differential recruitment of regions by the control group, suggesting more selective processing determined by the text properties. Additionally, the autism group had lower functional connectivity within the Theory of Mind network (between the left medial frontal and right superior temporal areas) as well as between the frontal portion of the Theory of Mind network (the left medial frontal gyrus) and the classical language processing regions (left inferior frontal and left middle/superior temporal regions) during the processing of intentional inferences. Morphometric measures indicated that that the autism group was typified by a smaller corpus callosum size than the control group, and furthermore, that the size of the relevant corpus callosum segment was correlated with functional connectivity.

2.2. Distribution of activation

The participants with autism recruited more right hemisphere regions, such as the right middle and superior temporal gyrus, than control participants (as shown in Fig. 2). This effect was similar across all three types of inferences (physical, emotional, and intentional). This higher right hemisphere activation in autism is probably due to the greater difficulty this group has in making inferences during discourse processing. Several studies have shown that people with autism have difficulty constructing a situation model of discourse (Happe, 1994; Minshew et al., 1995). This is especially important considering their difficulty in inferring intentions (Happe, 1994, 1995; Jolliffe & Baron-Cohen, 1999). The integration of information across sentences that is required for building an understanding of the speaker’s communicative intent was difficult for the participants with autism, as indicated by the greater engagement of the right hemisphere (compared to the control group) across all types of inferences.

Another striking difference between the autism and control groups was in the differential recruitment of regions for the different types of inferences. For all three types of inferences, the autism group activated a predominantly bilateral network (left inferior frontal gyrus, bilateral middle and superior temporal gyrus, bilateral angular gyrus, medial frontal gyrus and bilateral intraparietal sulcus) (as shown in Fig. 2). By contrast, the activation of the control group differentiated among the inference types, exhibiting more active mediation in medial frontal gyrus and bilateral temporo-parietal junction while processing intentional inferences than for physical or emotional inferences. In other words, the participants with autism seemed to be drawing on several additional regions in all of the conditions to cope with the difficulty and accomplish the task regardless of inference type. The control group showed a more adaptive recruitment of regions. The entire list of activated regions for the three types of inferences in the two groups is shown in Supplemental Table 1.

To determine more precisely how the discourse networks were different between the groups, activation for the three inferences as well as the context sentences was directly contrasted across groups. For simplicity of presentation, the comparison of the activation during the processing of the context (first two) sentences was collapsed across the three passage types because there
was little difference in the activation associated with the processing of the context sentences across the three conditions in either group. The individuals with autism showed greater activation than controls predominantly in the right hemisphere, specifically in the right middle to superior temporal gyrus, extending into the angular and supramarginal gyrus to varying degrees across conditions, as shown in Fig. 3. The areas that showed more activation in autism for the context sentences were the same areas that showed more activation for the autism group in two of the inference conditions (physical and emotional). This finding implies that the discourse task in general placed additional demands on the language network for participants with autism and that the difficulty of the task was manifested as early as the context sentences. It is perhaps noteworthy that the contrast between the autism and control groups revealed little group difference in the activation in occipital and left hemisphere language areas.

While the majority of areas that were more active in the participants with autism than the control group were common across the three types of inferences, there were some small areas of activation in autism that were specific to each type of inference. Although none of these areas were reliably more activated in the direct contrasts across groups for each inference type, we mention them here for completeness. The intentional inferences resulted in activation in autism of an area of the right caudate that was not present in the other two types of inferences. The physical inferences resulted in a cluster of activation in autism on the border of the superior parietal, superior occipital and middle occipital lobe that could be part of the dorsal stream in visual processing. The emotional inferences resulted in two small additional clusters of activation in autism, one that spread across the left angular gyrus and left inferior parietal lobe and a second small cluster in the left hemisphere portion of the superior medial frontal lobe.

To further examine the right temporoparietal activation difference across inference types in the two groups of participants (shown in the green ellipses in Fig. 3), an additional exploratory analysis was performed. A direct contrast of the difference in Beta-weights (taken from the general linear model) of the peak RTPJ voxel for the emotional, physical and intentional inference regressors revealed that the emotional and physical inference conditions for the controls differed from the intentional inference condition (for the controls) as well as all three inference types for the individuals with autism. The mean of these two conditions in which there was no activation in this area (control’s emotional and physical inferences) was lower than the mean of the conditions in which there was activation (control’s intentional, and all three inference types for the individuals with autism); this contrast was significant ($F_{(1,102)} = 5.30$, MSe = 0.00003742, $P = 0.02$).

2.3. Functional connectivity

The functional connectivity was systematically lower in participants with autism than in the control participants in predicted pairs of regions. More specifically, two central findings of functional connectivity analysis are: (a) the participants with autism showed reliably lower functional connectivity within the Theory of Mind network (between left medial frontal gyrus and right temporoparietal junction) than the control participants, and (b) the participants with autism showed lower functional connectivity between language and Theory of Mind processing regions than the control participants. Thus, there was evidence...
Fig. 3. The right temporal and right inferior frontal spillover processing (blue ellipses), right temporo-parietal junction and medial frontal Theory of Mind processing (green ellipses) that were more active in the autism group than the control. Activation is projected onto the surface rendering. For this contrast there were virtually no activated clusters in the left hemisphere.

of within and between network reductions in functional connectivity.

One of the ROIs that prominently entered into the autism underconnectivity was the left medial frontal gyrus. This ROI had lower functional connectivity in autism to the right temporo-parietal junction as well as to several portions of the left hemisphere language network (a full description of the networks appears in the experimental procedures). The set of six pairs of underconnected ROIs involving the left medial frontal gyrus are shown in Table 2. In the network level analysis (where the mean functional connectivity is computed for all of the ROI pairs in a network), the participants with autism had lower functional connectivity (0.53) between the left hemisphere language network (left inferior frontal gyrus and posterior left middle temporal gyrus) and the Theory of Mind network (left medial frontal gyrus and right temporo-parietal junction) in the intentional inference condition than did the control group (0.64); this effect was significant ($t_{18} = 1.74$, $P < 0.05$ one-tailed). The autism group also had lower functional connectivity between the left medial frontal ROI and both the left hemisphere language network and the text integration network (left frontal–left anterior temporal) in the intentional inference condition. (This was the only condition in which the control group showed activation in the right temporo-parietal junction portion of the Theory of Mind network, allowing calculation of functional connectivity). The analyses of the functional connectivity for the Theory of Mind networks and components were treated as planned comparisons and were not corrected for multiple contrasts. These network-level functional connectivity results are shown at the bottom of Table 2. In summary, when the control group showed activation in the right temporo-parietal junction portion of the Theory of Mind network, it did so with a greater connectivity between the left medial frontal portion of the Theory of Mind network to several language related individual ROIs as well as the language-related networks.

2.4. Corpus callosum size

The corpus callosum as a whole was smaller in individuals with autism (ratio to total brain volume = 0.584; raw size = 647 mm), than in controls (ratio = 0.646; raw size = 715 mm). Furthermore, the anterior portion was significantly smaller in the individuals with autism (autism: ratio = 0.293; raw size = 325; control: ratio = 0.327; raw size = 362 mm). When this region was further subdivided, the genu alone was significantly smaller in the group with autism (ratio = 0.109; raw size = 121 mm) than in the control group (ratio = 0.127; raw size = 141 mm). These differences held even when the measurement of the corpus callosum was left in raw measurements and not normalized to the size of the individual’s brain. It is believed that genu fibers can be mapped to prefrontal cortical areas (Witelson, 1989). Our results are in good agreement with the literature (Courchesne, Press, & Yeung-Courchesne, 1993; Egaas, Courchesne, & Saitoh, 1995; Hardan, Minshew, & Keshavan, 2000; Manes et al., 1999; Piven, Bailey, Ranson, & Arndt, 1997; Saitoh, Courchesne, Egaas, Lincoln, & Schreibman, 1995), and they were obtained in the context of lowered functional connectivity in the autism group.

2.5. Corpus callosum size and functional connectivity

The relationship between functional connectivity between ROIs and corpus callosum size was compared within the autism group. Only the corpus callosum regions that showed a reliable group difference (the genu, the anterior most portion, and the total corpus callosum) were considered in the analyses. For the intentional inferences, functional connectivity between the
Connectivity between the pairs is expressed as a z-score; the t-value corresponds to the difference in connectivity between the two groups.

left medial frontal gyrus and the right temporo-parietal junction was positively correlated with size of the corpus callosum ($r = 0.72$). This is particularly important because this network is considered to be part of the ToM network. Within the physical passages, functional connectivity between the left medial frontal gyrus and the right middle temporal gyrus was positively correlated with size of the genu ($r = 0.78$) and the anterior portion of the corpus callosum ($r = 0.77$ for anterior right middle temporal and $r = 0.71$ for posterior right middle temporal). These correlations were high (above 0.7) and statistically significant (d.f. = 15, cutoff 0.70, $P < 0.05$ corrected for multiple comparisons).

### 3. Discussion

At a neural level, the group with autism processed extended discourse differently than age and IQ-matched controls. Overall, they demonstrated more right hemisphere activation when reading text than the controls. Moreover, this extra RH activation in autism was not adaptive to the type of processing required by the text; it occurred during all three types of inference conditions and even during the reading of the context sentences. In contrast, the control group activated the right hemisphere only under certain conditions. Therefore, the difficulty individuals with autism have with inferencing cannot be due to a failure to engage right hemisphere processing resources, as others have argued based on the results of behavioral studies with individuals with right hemisphere brain damage (Happé, Brownell, & Winner, 1999). Engagement of the right hemisphere by the autism group during text comprehension clearly occurs; however, the RH in autism may be performing somewhat different functions than in the control group. In autism, the RH activation during text comprehension may involve processes that attempt to deal with the added difficulty of comprehending complex language. By contrast, in controls, the RH activation during text comprehension may draw on specialized right hemisphere social and language processing resources that enhance the processing of the information.

#### 3.1. A “spillover” account of right temporal and inferior frontal activation during text comprehension in autism

The autism group had right temporal activation when reading the inference sentences regardless of the type of inference demanded by the text. Additional right inferior frontal activation occurred in all conditions except the physical inference sentences (the relevant activation is highlighted by the blue ellipses in Fig. 3). The additional right hemisphere activation for the autism group suggests that the LH language areas may have been sufficiently taxed during text comprehension so as to recruit these RH language homolog areas. This additional right hemisphere activation also occurred for the context sentences, indicating that understanding discourse was more demanding for the autism group irrespective of the type of discourse. This pattern suggests that a more general phenomenon, namely a spillover of cortical processing to less specialized ancillary areas, was responsible for the increased right hemisphere activation in the autism group. Moreover, even though some of the right temporal activation in autism occurred in regions that are proximal and slightly anterior to the foci of the typical Theory of Mind activation reported in other studies (e.g., Castelli et al., 2002; Pelprey, Morris, Michelich, Allison, & McCarthy, 2005; Saxe & Wexler, 2005), this activation in autism (highlighted by the green ellipses in Fig. 3) may nevertheless have been generated by spillover rather than ToM processing.

It might be tempting to attribute the increased right hemisphere activation in the autism group to coarse coding processes that activate widespread semantic interpretations and associations (Beeman, 1998; Virtue, Haberman, Clancy, Parrish, & Jung-Beeman, 2006). However, the RH activation occurred in the context sentences as well as the inference sentences, which should not have made any additional demands on coarse coding. Moreover, additional RH activation occurred not only in the areas associated with coarse coding (right temporal) but also in the right inferior frontal region and the medial frontal area, suggesting the involvement of processes not related to coarse coding. Thus, the spillover interpretation of the extra RH activation in autism provides a better account.

#### 3.2. Disruption of the Protagonist monitoring/Theory of Mind network: temporo-parietal junction and medial frontal gyrus

The activation pattern across sentences and conditions is distinctly different in individuals with autism than the controls in the right temporo-parietal junction and the medial frontal...
gyrus. Combined, these two regions play a key role in discourse processing. This network has been found to be activated in Theory of Mind tasks (Castelli et al., 2002; Gallagher & Frith, 2003; Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Martin & Weisberg, 2003; Moll, de Oliveira-Souza, Bramati, & Grafman, 2002), as well as discourse comprehension tasks (Bottini et al., 1994; Eviatar & Just, 2006; Ferstl & von Cramon, 2001, 2002; Ferstl, Rinck, & von Cramon, 2005; Nichelli et al., 1995; Xu, Kemeny, Park, Frattali, & Braun, 2005). Mason and Just (2006) attribute activation in this network to Protagonist monitoring.

The individuals with autism activate this Theory of Mind/Protagonist monitoring network but do so indiscriminately (i.e. in all conditions), suggesting that its functioning is impaired. Specifically, the individuals with autism had greater activation of the right temporo-parietal portion of the Theory of Mind network compared to controls in the physical and emotional inferencing conditions but not in the intentional inference condition; this is highlighted by the green ellipses in Fig. 3. This counterintuitive result is readily interpreted by examining the network activation for both groups as compared to the fixation condition, as shown in Fig. 2. The individuals with autism indiscriminately activated both the right temporo-parietal junction and medial frontal areas in all three conditions. In contrast, the control group selectively activated the right temporo-parietal junction (a ToM area) only in the intentional inference condition. Thus, the group subtraction (autism-control) shown in Fig. 3 indicates no group difference in the right temporo-parietal area in the intentional condition when activation occurred in this area for both groups.

The pattern of differential activation of the right temporal parietal junction can readily be seen in the contrasts of each condition against fixation for the two groups. Even though there may be some below threshold activation in the control group for the physical and emotional inferences, it is clear that this region is reliably activated only for the intentional inferences, indicating that, in the control group, it is fully recruited only when necessary.

In contrast, the activation of the medial frontal region is not restricted to the inference sentence but occurs throughout the processing of the text for the individuals with autism. In fact, the medial frontal region is more active in the autism group than the control group for the context sentences; however, during the inference sentences, there is no difference between the two groups with respect to this medial frontal region. This differential pattern across the two components of this network indicates that they have distinguishable functions. The right temporo-parietal junction may be involved with reasoning about particular mental states (Saxe & Kanwisher, 2003), whereas the medial frontal gyrus may more generally flag and keep track of possible intention-related inferences to be made. These proposed functions are consistent with the general view of the medial frontal gyrus as an executive processor and the description of the involvement of these regions in Theory of Mind tasks (Gallagher & Frith, 2003). Thus, the medial frontal region of the Protagonist monitoring network may increase in activation whenever an inference is to be made, but the posterior portion of the network may become activated only when the protagonist’s intention (ToM) is a component of that inference. The temporo-parietal activation for intentional inferences could be related to reasoning about a protagonist’s intention based on one’s own experience.

### 3.3. Inefficiency in the Theory of Mind network in autism

The functional connectivity results for the autism group furthermore suggest an inefficiency of communication between key components of the Theory of Mind network. The functional connectivity analysis yielded three main results indicating that inefficiency in the Theory of Mind network constitutes part of the autism group’s difficulty with discourse processing. First, the participants with autism had reduced functional connectivity, relative to controls, within the Theory of Mind network (between left medial frontal gyrus and right temporo-parietal junction). Second, the autism group showed reduced functional connectivity between Theory of Mind areas (left medial frontal gyrus) and the left hemisphere language network (left inferior frontal gyrus, and left superior temporal gyrus). Third, the functional connectivity between frontal and parietal regions was positively correlated with the size of the anterior-most subregion of the corpus callosum (genu) in autism. Therefore, the coordination within as well as between critical networks was found to be reduced in autism and the functional measures were correlated with anatomical differences in the anterior portion of the corpus callosum through which the fibers that connect the two medial frontal regions would be expected to pass. Thus, there is a reduced ability for the left medial frontal region to be connected to the right temporal parietal region; this having been done through an indirect link via the right medial frontal region. Additionally, the genu of the corpus callosum may contain fibers by which the frontal and parietal regions cross (de Lacoste, Kirkpatrick, & Ross, 1985). This pattern of results suggests that the key regions involved in accomplishing the inferential processing tasks were less coordinated in autism and that this underconnectivity was more pronounced in connections between frontal and parietal regions. This underconnectivity in autism emerged despite higher autism group activation in these regions. Underconnectivity in critical brain networks, particularly frontal and parietal regions, has been found in several other studies of individuals with autism. These networks include language comprehension (Just et al., 2004), language and working memory (Koshino et al., 2005), language and visual imagery (Kana et al., 2006), executive functioning (Just et al., 2007), and Theory of Mind (Castelli et al., 2002).

The results of this study should not be interpreted as supporting the view that autism is primarily a deficit in Theory of Mind. Although there clearly is such a deficit, it may be secondary to a more general underlying neurofunctional basis of autism. That is, the individuals with autism had a less integrated Theory of Mind network not because of a deficit in that network per se but because the extensive network integration demands were beyond the capacity of the functional and structural capabilities of the cortical system. Thus, the Theory of Mind network fails to
operate efficiently, but the processing deficit is not restricted to that network. This view is supported by the results of other studies that demonstrated the failure of the autism group to recruit efficient processing systems in tasks that did not involve the Theory of Mind network (Just et al., 2004, 2007; Kana et al., 2006; Koshino et al., 2005).

4. Summary

The overall pattern of activation, that is, increased right temporal and right inferior frontal activation, indicates that general aspects of language comprehension as well as specific aspects of comprehending a narrative text (monitoring a protagonist’s state of mind) were effortful tasks for the autism group. These results are consistent with a spillover account, in which the RH homologs are recruited when the LH language areas are taxed. In addition, the greater activation of the right temporo-parietal junction and the medial frontal areas of the autism group in various contrasts, as well as decreased functional connectivity, indicate an inefficient or disrupted Theory of Mind network in autism.

In autism, text processing presents a challenge, probably for the same reasons that understanding life in general is difficult. Understanding the actions and intentions of others is part of the challenge, to be sure, but so is the integration of a myriad of facts about the world that have to be inferred and integrated during discourse comprehension. Whenever such discourse challenges arise, the cortical system in autism attempts to meet the challenge by engaging RH areas indiscriminately. However, network connectivity (functional and structural) limits how effectively the recruited cortical networks can function. More specifically, the altered processing of text by the individuals with autism supports the conclusion that the Theory of Mind network, though activated in individuals with autism, is inefficient.

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, MH029617 from the National Institute of Mental Health Grant, and K23 DC006691 from the National Institute of Deafness and Other Communication Disorders. Address correspondence to: Robert Mason, Center for Cognitive Brain Imaging, Carnegie Mellon University, Department of Psychology, Pittsburgh, PA 15213 or email: rmason@andrew.cmu.edu. We appreciate the assistance of current and former members of the Center for help in conducting the experiments. In particular, we would like to thank Sarah Schipul for her assistance in the functional connectivity analysis, Sarah Berson and Vanessa Gorley for their assistance in measuring the corpus callosum, Stacey Becker for editorial comments, Kelsey Woods for developing additional materials, and Chantel Prat for comments on an earlier version of the manuscript. We would also like to express our sincere appreciation for the time and effort of the participants and their families in making this research possible.

Appendix A. Supplementary data


References


