

Published in final edited form as:

Soc Neurosci. 2011 ; 6(0): 569–581. doi:10.1080/17470919.2011.620774.

Theory of mind in schizophrenia: Exploring neural mechanisms of belief attribution

Junghee Lee^{1,2}, Javier Quintana^{1,2}, Poorang Nori^{1,2}, and Michael F. Green^{1,2}

¹Semel Institute for Neuroscience and Human Behavior, UCLA

²VA Greater Los Angeles Healthcare System

Abstract

Background—Although previous behavioral studies have shown that schizophrenia patients have impaired theory of mind (ToM), the neural mechanisms associated with this impairment are poorly understood. This study aimed to identify the neural mechanisms of ToM in schizophrenia using functional magnetic resonance imaging (fMRI) with a Belief Attribution Task.

Methods—In the scanner, 12 schizophrenia patients and 13 healthy control subjects performed the Belief Attribution Task with 3 conditions: a false belief condition, a false photograph condition, and a simple reading condition.

Results—For the false belief vs. simple reading conditions, schizophrenia patients showed reduced neural activation in areas including the temporo-parietal junction (TPJ) and medial prefrontal cortex (MPFC) compared with controls. Further, during the false belief vs. false photograph conditions we observed increased activations in the TPJ and the MPFC in healthy controls, but not in schizophrenia patients. For the false photograph vs. simple reading condition, both groups showed comparable neural activations.

Conclusions—Schizophrenia patients showed reduced task-related activation in the TPJ and the MPFC during the false belief condition compared with controls, but not for the false photograph condition. This pattern suggests that reduced activation in these regions is associated with, and specific to, impaired ToM in schizophrenia.

Keywords

theory of mind; belief attribution; schizophrenia; fMRI; social cognition

INTRODUCTION

Schizophrenia is a complex and severe mental disorder that affects approximately 1% of the population worldwide. It is associated with poor functioning that is observed in the form of poor family relationships, difficulty in maintaining employment, and social withdrawal. Impaired functioning in daily life is present before the onset of psychosis (Davidson et al., 1999) and tends to persist throughout the course of illness (Murray & Lopez, 1997; WHO, 2008). Psychotic symptoms are not generally strongly associated with functional outcome, but other factors, such as cognition and negative symptoms (e.g. lack of drive, flat affect), appear to be consistent determinants (M. F. Green, 1996; M. F. Green, Kern, Braff, & Mintz, 2000; Milev, Ho, Arndt, & Andreasen, 2005). Within the general area of cognition,

Correspondence: Junghee Lee, PhD, 300 Medical Plaza Rm 2261, Semel Institute for Neuroscience and Human Behavior, UCLA, Los Angeles, CA 90095-6968, jungheele@ucla.edu.

Financial disclosure: None of the authors has any to disclose.

social cognition has been shown to be a key determinant of poor functioning in schizophrenia (Kerr & Neale, 1993; Penn, Corrigan, Bentall, Racenstein, & Newman, 1997; Penn, Ritchie, Francis, Combs, & Martin, 2002). This study aims to investigate the neural mechanisms in schizophrenia of one area of social cognition, theory of mind, using functional magnetic resonance imaging (fMRI).

Theory of mind (ToM; also called mental state attribution) is a social cognitive construct that refers to the ability to make high-level inferences about one's own and other persons' mental states such as thoughts, beliefs, desires and feelings (Premack & Woodruff, 1978). Such an ability starts at a very early stage of development (around ages 3–4) (Dennett, 1978; Leslie, 1987). From an operational perspective, ToM makes it possible to assess or understand someone else's mental states in specific situations, and thus to interpret and anticipate their behaviors (e.g. He is missing his wallet. He is going back to his office because he thinks he left his wallet there). The ability to understand and predict others' behavior has obvious relevance for successful social interactions.

Previous studies using a variety of behavioral paradigms have found that schizophrenia patients show impairment in their ability to attribute mental states to others. Schizophrenia patients showed poor performance on picture-sequencing tasks (Sarfati, Hardy-Bayle, Besche, & Widlocher, 1997), on tasks detecting irony or sarcasm (Kern et al., 2009; Mitchley, Barber, Gray, Brooks, & Livingston, 1998), on tasks attributing spontaneous mental states to non-human objects (Horan et al., 2009), and on tasks of false-belief stories (Frith & Corcoran, 1996; Pickup & Frith, 2001). A recent meta-analysis (Bora, Yucel, & Pantelis, 2009) found a large effect size (1.10) for the patient - control difference and the effect size remained large when the comparison was focused on remitted patients only (.80). Further, the patient-group difference appears to be similar across all phases of the illness. A recent study from our group (M.F. Green et al., in press) compared the patient-group differences in ToM among individuals considered to be in a prodromal phase, first-episode schizophrenia patients and chronic schizophrenia patients and found that the group differences were relatively constant across all phases of illness, indicating good stability of the social cognitive impairment.

One might wonder whether impaired ToM in schizophrenia is due to general cognitive impairments; however, studies have shown that impaired ToM in schizophrenia cannot be explained by non-social cognitive impairments (Langdon, Coltheart, Ward, & Catts, 2001; Mazza, De Risio, Surian, Roncone, & Casacchia, 2001; Schenkel, Spaulding, & Silverstein, 2005). For example, one study (Schenkel, et al., 2005) showed that ToM in schizophrenia was not significantly related to verbal fluency or verbal intelligence. Another study (2001) failed to find a significant correlation between ToM and executive function (measured by Wisconsin Card Sorting Test) or general intelligence. In addition, Langdon et al. (2001) showed that schizophrenia patients had performance comparable to controls when making inferences about non-social cause-and-effect relations or social knowledge in general (without ToM component), but they showed impaired ToM performance. These findings suggest that, even though schizophrenia patients have impaired cognitive functions, impaired ToM in schizophrenia cannot be fully explained by deficits in other cognitive processes.

Most of the studies of neural mechanisms associated with ToM in healthy individuals employed either a false belief task or a cartoon task. Commonly activated areas include the medial prefrontal cortex, the temporal-parietal junction (TPJ), and the precuneus/posterior cingulate cortex (Frith & Frith, 2006; Gallagher & Frith, 2003; Saxe & Kanwisher, 2003; Saxe & Powell, 2006; Siegal & Varley, 2002), known as ToM network. A general question is whether activation of these areas is due to the general reasoning demands of the task or is

specific to attributing mental states to other people. Kanwisher, Saxe and colleagues attempted to address this issue by conducting a series of fMRI studies that compared regional brain activation during stories that involved belief attribution with stories that required similar reasoning ability, but without belief attribution (Saxe & Kanwisher, 2003; Saxe & Powell, 2006). For example, they compared the neural activations of belief attribution with non-social reasoning ability. To measure belief attribution, false belief stories were used in which subjects were asked to make inferences about the belief of another person that differs from the current state. To measure non-social reasoning ability, a false photograph task was used that required subjects to make inferences about the physical world that differs from the current state of the world. In other words, both the false belief task and the false photograph task asked subjects to inhibit the current or “true” representation and make a decision based on the “false” representation of the belief state (as in the false belief task) or the physical world (as in the false photograph task). While both the false belief task and the false photograph task requires inhibitory control process, only the false belief condition involves belief attribution, one component of ToM. They found that subjects showed greater activation in the medial prefrontal cortex, the bilateral TPJ and the posterior cingulate cortex when they read stories about belief attribution than when they read the false photograph stories. In a subsequent study, Kanwisher and Saxe added a human non-belief condition and showed that activation in those brain regions was not due to the fact that stories involved human beings; instead, were specific to the belief attribution component (Saxe & Kanwisher, 2003; Saxe & Powell, 2006). It is possible that these brain regions are only related to the belief attribution component of ToM but not to ToM in general (e.g. attributing emotion, desire, or other mental states). Several studies showed that the bilateral TPJ and the posterior cingulate cortex were activated only during the belief attribution, whereas the medial prefrontal cortex was activated for belief attribution and also other mental state attributions (Saxe & Powell, 2006** other refs). These findings suggest that the TPJ and posterior cingulate cortex may be specifically involved in belief attribution but the medial prefrontal cortex is related to ToM more broadly.

Only a few studies have examined ToM in schizophrenia with functional neuroimaging. One fMRI study (Russell et al., 2000) reported reduced activation in the left prefrontal cortex when schizophrenia patients were asked to describe the mental state reflected in photographs of eyes. Another study used a cartoon task to evaluate how people infer social intention and found decreased activations in the TPJ and the medial prefrontal cortex in schizophrenia patients (Walter et al., 2009). In contrast, another study that used a similar cartoon task found increased activations in the medial prefrontal cortex and TPJ in schizophrenia patients (Brune et al., 2008). Although these studies suggest that abnormal brain activation is associated with impaired ToM in schizophrenia, the directions of the findings are inconsistent across studies. Further, the methods are limited to visual assessment of ToM, and they did not fully control for the extent to which non-specific task-related demands.

In this study we investigated the neural mechanisms of ToM in schizophrenia using a well-validated fMRI paradigm in social neuroscience: the Belief Attribution Task (Saxe & Kanwisher, 2003). This task, as adapted for use in schizophrenia, consists of 3 conditions: a false belief condition, a false photograph condition and a simple reading condition. In the false belief condition, a character’s belief is false; that is, it is different from the actual situation (Dennett, 1978). To perform this condition well, subjects must predict the character’s behavior based on the character’s (inaccurate) belief, not based on the actual situation. The false photograph condition was intended to control for the general problem-solving structure of the false belief vignette (Zaitchik, 1990). These vignettes have the same story structure but differ in that the false photograph condition requires the similar general reasoning ability without the belief attribution. To adapt this task for use in schizophrenia, we added a reading condition as a control for the reading ability and associated regional

activity required by the false belief condition. Thus, having three conditions allowed us to determine neural substrates of ToM deficits in schizophrenia while controlling for other non-specific task-related demands (i.e. general reasoning ability, ability to read simple vignettes). With the Belief Attribution Task, we hypothesized that schizophrenia patients would show reduced task-related activations in brain regions associated with ToM during the false belief condition, but not during the false photograph condition, compared to healthy controls.

METHODS

Participants

Fourteen (3 female) patients with schizophrenia and 14 (3 female) healthy controls participated in this study. Schizophrenia patients were recruited from outpatient clinics at the Veterans Affairs Greater Los Angeles Healthcare System (VAGLAHS) and from local board and care facilities. Patients met diagnostic criteria for schizophrenia according to the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID) (First, Spitzer, Gibbon, & Williams, 1997). Exclusion criteria for patients included: 1) substance abuse or dependence in the last six months, 2) mental retardation based on review of medical records, 3) history of loss of consciousness for more than one hour, 4) an identifiable neurological disorder, or 5) insufficient fluency in English to understand testing procedures. All patients were medicated and clinically stable at the time of testing.

Healthy control participants were recruited through flyers distributed in the local community and website postings. Exclusion criteria for control participants included: 1) history of schizophrenia or other psychotic disorder, bipolar disorder, recurrent depression, history of substance dependence, or any substance abuse in the last 6 months based on the SCID (First, et al., 1997), 2) any of the following Axis II disorders: avoidant, paranoid, schizoid, or schizotypal, based on the SCID for Axis II disorders (First, Gibbon, Spitzer, Williams, & Benjamin, 1996), 3) schizophrenia or other psychotic disorder in a first-degree relative, 4) any significant neurological disorder or head injury of loss of consciousness for more than one hour, or 5) insufficient fluency in English to understand testing procedures.

All SCID interviewers were trained to a minimum kappa of 0.75 for key psychotic and mood items through the Treatment Unit of the VA VISN 22 Mental Illness Research, Education, and Clinical Center (MIRECC). All participants were evaluated for the capacity to give informed consent and provided written informed consent after all procedures were fully explained, according to procedures approved by the Institutional Review Boards at UCLA and the VAGLAHS.

Design and Procedure

The Belief Attribution Task was modeled after Saxe and Kanwisher (2003) and was composed of three conditions: a false belief condition, a false photograph condition, and a simple reading condition (see examples in Table 1). In the false belief condition, subjects were presented with vignettes in which they needed to infer the beliefs of a character, even when these beliefs were different from the actual state of affairs. The false photograph vignettes had the same story structure and required the same level of complex reasoning as the false belief condition, but lacked the belief attribution component. The simple reading condition controlled for the process of reading the false belief condition required, and consisted of stories describing non-human objects.

All subjects were presented with 12 vignettes (average number of words = 32) for each condition, and each vignette was accompanied by a single two-alternative forced choice “fill-in-the-blank” question. The fill-in-the-blank question consisted of a single sentence

with a word missing, presented above two alternative completions. For the false belief and false photograph conditions, 50% of the questions asked about the content of false representation; the other 50% asked about the actual outcome of the story. For the reading condition, all of the questions were related to description of non-human objects.

At the onset of each trial, a vignette was presented for 12 seconds; next a fill-in-the-blank question was presented for 10 seconds while the vignette was still visible. After the vignette and question disappeared, a probe was presented for 3 s, prompting the subject's response. Subjects responded by pressing the corresponding button with their dominant hand. Then, an intertrial interval (ITI) that was jittered between 12 and 18 seconds ensued. The Belief Attribution Task consisted of 6 runs, each lasting 4 minutes and 28 seconds, with 6 trials per run (2 trials of each condition). All tasks were presented through MR-compatible LCD goggles (Resonance Technology, Northridge, CA). A schematic diagram of the procedure is shown in Figure 1.

fMRI data acquisition

All scanning was conducted on a 3T scanner (Siemens Trio, Erlangen, Germany) located in the UCLA Ahmanson Lovelace Brain Mapping Center. For anatomical reference, a high-resolution echo planar axial T2-weighted series was obtained for each subject prior to functional scanning (TR = 5000 ms, TE = 30 ms, flip angle = 90 degrees, 33 slices, FOV 22 cm). A T2*-weighted gradient-echo sequence was used to detect blood-oxygen level-dependent (BOLD) signal (TR=2000 ms, TE=30 ms, flip angle=75 degrees, voxel size of $3.4 \times 3.4 \times 4.00$ mm), acquiring 33 slices parallel to the AC-PC plane.

fMRI data analysis

Imaging data were analyzed using the FEAT (FMRI Expert Analysis Tool) version 5.98, part of FMRIB Software Library (FSL, Smith et al., 2004). The pre-statistics processing included motion correction using MCFLIRT (Motion Correction using FMRIB's Linear Imaging Registration Tool) (M. Jenkinson, Bannister, Brady, & Smith, 2002), non-brain removal using BET (Brain Extraction Tool) (Smith, 2002) spatial smoothing using a Gaussian kernel of the full width at half maximum (FWHM) 5 mm, grand-mean intensity normalization by a single multiplicative factor, and high pass temporal filtering (Gaussian weighted LSF straight line fitting with $\sigma = 50.0$ s). To facilitate multi-subject analyses, statistical images created for each subject were normalized into a standard space from the Montreal Neurological Institute (MNI) using affine transformation with FLIRT (FMRIB's Linear Image Registration Tool) (M. Jenkinson & Smith, 2001)(Jenkinson and Smith, 2001).

Functional images were analyzed using FLIM (FMRIB's Improved Linear Model) with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith, 2001). First, for each run and each subject, data from all 3 conditions were modeled by convolving them with a canonical hemodynamic response function and temporal derivatives were included as covariates of no interest to increase statistical sensitivity. We computed 3 main contrasts of interest for each run and each subject: false belief vs. simple reading, false photograph vs. simple reading, and false belief vs. false photograph. Second, to average across the 6 runs, we completed a second-level analysis using a fixed-effects model, by forcing the random effect variance to zero in FLAME (FMRIB's Local Analysis of Mixed Effects) (Beckmann, Jenkinson, & Smith, 2003; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004). Third, to characterize functional activations in each group separately and to directly compare activations of patients to those of controls for each contrast of interest, a mixed-effects model (FLAME stage 1+2) (Beckmann, et al., 2003; Woolrich, et al., 2004) was performed and the resulting statistical images were thresholded using a z value > 2.3 and a cluster

probability of $p = 0.05$, corrected for whole-brain multiple comparisons using Gaussian random field theory (Worsley, 2001) unless otherwise noted.

RESULTS

Two patients were excluded from analyses due to below-chance level performance during the Belief Attribution Task (below 50% accuracy across the 3 conditions), and one control was excluded due to technical problems during scanning. Therefore, 12 schizophrenia patients (2 female) and 13 Healthy controls (3 females) were included in the following analyses.

Demographic Information and Behavioral Performance

Table 2 presents demographic information and behavioral performance during the Belief Attribution Task. Schizophrenia patients and healthy controls were comparable in terms of age, parental education, and gender (age, $t_{23} = -1.1$, $p = 0.27$; parental education, $t_{23} = -1.09$, $p = 0.28$; gender, $\chi^2 = 1.60$, $p = 0.68$), but not in their own education levels ($t_{23} = -2.19$, $p < 0.05$).

Behavioral performance of both groups during the Belief Attribution Task was examined using a 3 (condition) by 2 (group) repeated measures ANOVA with condition as a within-subject factor and group as a between-subject factor. We found a significant main effect of condition ($F_{2,46} = 24.32$, $p < 0.001$), and a significant main effect of group ($F_{1,23} = 16.31$, $p < .01$). The condition by group interaction was not significant. Schizophrenia patients performed worse than controls across all conditions. Across groups, performance was best on the simple reading condition, intermediate on the false belief condition, and worst on the false photograph condition. Each condition was significantly different (p 's $< .05$) from the other after correction for multiple comparisons.

fMRI activations during the Belief Attribution Task

Table 3 lists brain areas (local maxima of the clusters) that exhibited above-threshold activations for each contrast of interest in the healthy control and schizophrenia patient groups. For the false belief vs. simple reading contrast (see Figure 2), healthy controls showed increased activations in several areas, including the bilateral TPJ, bilateral middle temporal gyri, bilateral angular gyri, precuneus, right middle frontal gyrus, and medial prefrontal cortex. Schizophrenia patients showed increased activations in the right TPJ, precuneus and right supramarginal gyrus. When the groups were compared directly to each other, controls showed significantly more activation relative to patients on the bilateral TPJ, left middle temporal gyri, right medial prefrontal cortex, putamen, globus pallidus, and right amygdala.

For the contrast of the false photograph versus simple reading conditions, controls showed activations in the bilateral TPJ, right supramarginal gyrus and left middle temporal gyrus; but patients only activated the anterior portion of superior frontal gyrus and middle frontal gyrus. However, a direct group comparison did not reveal any brain regions significantly different between groups above the threshold.

Finally, we examined the contrast between the false belief and false photograph conditions (Figure 3). Healthy controls activated the bilateral TPJ, bilateral angular gyri, middle temporal gyrus, medial prefrontal cortex, posterior cingulate gyrus, and precuneus. Schizophrenia patients, on the other hand, only showed increased activations in the precuneus and posterior cingulate cortex. Direct group comparison revealed that relative to patients, controls showed significantly increased activations in several brain regions, including the medial prefrontal cortex, posterior cingulate cortex, precuneus, and parahippocampus.

DISCUSSION

This study examined neural correlates of ToM in schizophrenia. Using three conditions from the Belief Attribution Task (false belief, false photograph, and simple reading) we evaluated neural activation that was specific to belief attribution while controlling for neural processes associated with general reasoning or reading ability. For the contrast between the false belief and simple reading conditions, schizophrenia patients exhibited significantly less activation compared to controls in several brain regions including the bilateral TPJ and right medial prefrontal cortex. However, both schizophrenia patients and healthy controls showed comparable patterns of neural activations for the contrast between the false photograph, which required reasoning similar to the false belief condition but without mental state attribution, and the simple reading conditions. Finally, when comparing the false belief and false photograph conditions, we observed significantly less activation of schizophrenia patients relative to controls in several regions, including the medial prefrontal cortex, anterior cingulate cortex and precuneus. Our findings show reduced task-related neural activations in schizophrenia in several brain regions that have been associated with ToM tasks in healthy controls.

The patients showed lower accuracy than controls on the false belief condition. This performance difference, however, is unlikely to explain the differential patterns of neural activation schizophrenia patients exhibited during the false belief condition. Performance differences were seen across all conditions and the group by condition interaction was not significant. In addition, both groups performed less accurately during the false photograph condition than during the false belief condition (i.e. both patients and controls found the former more difficult than the latter.) In spite of such behavioral findings, we observed greater fMRI activation differences between groups in the false belief condition than in the false photograph condition. Hence, reduced task-related neural activation in schizophrenia appears to be a specific abnormality in belief attribution as opposed to a result of other non-specific task-related factors. The current findings suggest that reduced neural activations in key regions of ToM network is associated with impaired ToM in schizophrenia.

We observed that healthy controls showed increased activations in the TPJ bilaterally, medial prefrontal cortex and precuneus when belief attribution was compared with general reading ability, as well as when it was compared with non-social reasoning. Previous studies in healthy individuals showed the critical role of the right TPJ – and to a lesser extent left TPJ, medial prefrontal cortex and precuneus in belief attribution (Saxe & Powell, 2006; Saxe & Wexler, 2005). The current finding provides further supports to the critical role of these areas in ToM processes (Frith & Frith, 2006; Gallagher & Frith, 2003; Saxe & Kanwisher, 2003; Saxe & Powell, 2006; Siegal & Varley, 2002). Among these areas, schizophrenia patients showed increased activations in the right TPJ and precuneus when belief attribution was compared with reading ability. However, when belief attribution was compared with non-social reasoning, schizophrenia patients showed increased activation in precuneus only. The current findings of schizophrenia patients suggest that patients may not have a fully functionally specialized neural network for inferring mental state as controls do.

In the few previous reports of TPJ activity during ToM tasks between schizophrenia patients and controls findings have been inconsistent, including reports of hypoactivations of patients in bilateral TPJ (Walter, et al., 2009)(Walter et al, 2009), hyperactivations of patients in left TPJ (Brune, et al., 2008), and no group differences (Brunet, Sarfati, Hardy-Bayle, & Decety, 2003). It is possible that these inconsistent findings of TPJ activation in schizophrenia may arise because the ToM tasks that were used varied in the extent to which they were focused on belief as opposed to other types of mental states (e.g., intention, affective state, preference). Among the ToM network, the TPJ and precuneus have been associated more

narrowly with belief attribution (Saxe, Moran, Scholz, & Gabrieli, 2006; Saxe & Wexler, 2005). Hence, studies using ToM tasks focusing on belief attribution, such as ours, are likely to probe the ToM-related neural activations in the TPJ. Our findings suggest that the hypoactivation in the TPJ may underlie impaired belief attribution in schizophrenia.

It is also noteworthy that schizophrenia patients failed to show increased activations in the medial prefrontal cortex during the belief attribution, in contrast to controls. The medial prefrontal cortex has been associated with theory of mind beyond belief attribution (e.g., intent, affective states, preference, ambiguous mental state) (A. C. Jenkinson & Mitchell, 2010; van Overwalle, 2009). In addition, the medial prefrontal cortex has been also involved in other areas of social cognition in which schizophrenia patients showed impairments (Harvey, Lee, Horan, Ochsner, & Green, under revision; Sergi & Green, 2003), such as self-related information processing and integration of diverse social cues (Kelley et al., 2002; Mitchell, Macrae, & Banaji, 2006). It might be possible that the medial prefrontal cortex would also be dysfunctional during these social cognitive processes in schizophrenia. It will be of great importance to carefully examine the extent to which dysfunctional medial prefrontal cortex underlie social cognitive impairments in schizophrenia.

Recently there is growing debate on whether some areas of ToM network, especially the medial prefrontal cortex and TPJ, are exclusively involved with ToM or also involved with other cognitive processing such as attentional process or inhibitory control that are often necessary for ToM. For example, Mitchell (2008) showed overlapping activation in the right TPJ between ToM task and a selective attention task. A recent study by Rothmayr and colleagues (2011) also found that the medial prefrontal cortex and the right TPJ showed increased activation for both a ToM task and an inhibitory control task, whereas the left TPJ exclusively activated for a ToM task. In contrast, two recent studies (Scholz et al. 2009; Decety & Lamm, 2007) showed that, despite a small overlap between ToM and attentional activation tasks in the right TPJ, each task also activated neighboring, but separate, regions in the right TPJ. Finally, studies with brain damaged patients showed that patients with damage to the left TPJ showed impaired performance during the ToM task but not during a control task requiring general reasoning ability TPJ (Samson, Apperly, Chiavarino, & Humphreys, 2004). It appears that the so-call ToM network is indeed involved with ToM ability and it may be distinct from neighboring regions that are associated with attentional processing. Further studies with high-resolution functional neuroimaging techniques may be able to clarify the specific roles of sub-regions within ToM network.

In summary, we found that schizophrenia patients showed aberrant activation in the brain network associated with ToM when they inferred belief of another person, but not when they performed tasks that required similar reasoning demands but did not involve belief attribution component. Complementing a large number of other studies of ToM in schizophrenia that used performance task, this study demonstrates abnormal neural activations related to ToM processing in schizophrenia. Finally, the findings of the current study also open several promising avenues for future exploration. First, we did not assess cognitive abilities that may be related to performance on the false belief task in schizophrenia (e.g., executive function, inhibitory control, verbal intelligence). Although we included the false-photograph condition and a simple reading condition to control for non-social general reasoning ability and a basic reading ability, it would be helpful to include a wide range of neuropsychological assessments in future studies to elucidate the extent to which these abilities are related to belief attribution in schizophrenia. Second, this study focused on a specific component of ToM in schizophrenia, namely belief attribution. Third, we did not assess negative symptoms or community functioning of schizophrenia patients in this study. Deficits in ToM performance have been associated with poor community functioning of schizophrenia (Couture, Penn, & Roberts, 2006; Fett et al., 2011; Pijnenborg

et al., 2009; Roncone et al., 2002). Abnormal patterns of neural activation reported here provide a clearer view of the neural basis for social cognitive processes that are related to community functioning in schizophrenia. With larger subject samples, fMRI indices could be entered into statistical model to test their role as physiological determinants of impaired community functioning.

Acknowledgments

Funding / Support:

Support for this study came from Veterans Integrated Service Network 22 Mental Illness Research, Education and Clinical Center (to Lee); and Merit Review Award from the Veterans Administration Medical Research Service (to Quintana); and NIMH Grants MH043292 and MH065707 (to Green). We wish to thank Dr. Rebecca Saxe for providing the vignettes for the task.

For generous support of the UCLA Brain Mapping Center, we also thank the Brain Mapping Medical Research Organization, Brain Mapping Support Foundation, Pierson-Lovelace Foundation, The Ahmanson Foundation, William M. and Linda R. Dietel Philanthropic Fund at the Northern Piedmont Community Foundation, Tamkin Foundation, Jennifer Jones-Simon Foundation, Capital Group Companies Charitable Foundation, Robson Family, and Northstar Fund.

References

- Beckmann CF, Jenkinson M, Smith SM. General multilevel linear modeling for group analysis in fMRI. *Neuroimage*. 2003; 20:1052–1063. [PubMed: 14568475]
- Bora E, Yucel M, Pantelis C. Theory of mind impairment in schizophrenia: meta-analysis. *Schizophr Res*. 2009; 109(1–3):1–9. [PubMed: 19195844]
- Brune M, Lissek S, Fuchs N, Witthaus H, Peters S, Nicolas V, et al. An fMRI study of theory of mind in schizophrenia patients with “passivity” symptoms. *Neuropsychologia*. 2008; 46(7):1992–2001. [PubMed: 18329671]
- Brunet E, Sarfati Y, Hardy-Bayle MC, Decety J. Abnormalities of brain function during a nonverbal theory of mind task in schizophrenia. *Neuropsychologia*. 2003; 41(12):1574–1582. [PubMed: 12887982]
- Couture SM, Penn DL, Roberts DL. The functional significance of social cognition in schizophrenia: a review. *Schizophr Bull*. 2006; 32(Suppl 1):S44–63. [PubMed: 16916889]
- Davidson M, Reichenberg A, Rabinowitz J, Weiser M, Kaplan Z, Mark M. Behavioral and intellectual markers for schizophrenia in apparently healthy male adolescents. *Am J Psychiatry*. 1999; 156(9):1328–1335. [PubMed: 10484941]
- Dennett D. Beliefs about beliefs. *Behavioral and Brain Sciences*. 1978; 1:568–570.
- Fett AK, Viechtbauer W, Dominguez MD, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. [Research Support, Non-U.S. Gov't]. *Neuroscience and biobehavioral reviews*. 2011; 35(3):573–588. [PubMed: 20620163]
- First, MB.; Gibbon, M.; Spitzer, RL.; Williams, JBW.; Benjamin, L. *Structured Clinical Interview for DSM-IV Axis I Personality Disorders*. New York State Psychiatric Institute; New York, NY: 1996.
- First, MB.; Spitzer, RL.; Gibbon, M.; Williams, JBW. *Structured Clinical Interview for DSM-IV Axis I Disorders - Patient Edition*. New York State Psychiatric Institute; New York, NY: 1997.
- Frith CD, Corcoran R. Exploring ‘theory of mind’ in people with schizophrenia. *Psychol Med*. 1996; 26(3):521–530. [PubMed: 8733211]
- Frith CD, Frith U. How we predict what other people are going to do. *Brain Res*. 2006; 1079(1):36–46. [PubMed: 16513098]
- Gallagher HL, Frith CD. Functional imaging of ‘theory of mind’. *Trends Cogn Sci*. 2003; 7(2):77–83. [PubMed: 12584026]
- Green MF. What are the functional consequences of neurocognitive deficits in schizophrenia? *Am J Psychiatry*. 1996; 153(3):321–330. [PubMed: 8610818]

- Green MF, Bearden CE, Cannon TD, Fiske AP, Helleman G, Horan WP, et al. Social cognition across phases of illness in schizophrenia. *Schizophrenia bulletin*. in press.
- Green MF, Kern RS, Braff DL, Mintz J. Neurocognitive deficits and functional outcome in schizophrenia: are we measuring the “right stuff”? *Schizophr Bull*. 2000; 26(1):119–136. [PubMed: 10755673]
- Harvey PD, Lee J, Horan WP, Ochsner K, Green MF. Do patients with schizophrenia benefit from a self-referential memory bias? under revision.
- Horan WP, Green MF, Wynn JK, Lee J, Castelli F, Nuechterlein KH. Disturbances in the spontaneous attribution of social meaning in schizophrenia. *Psychological Medicine*. 2009; 39(4):679–687.
- Jenkinson AC, Mitchell JP. Mentalizing under uncertainty: Dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cereb Cortex*. 2010; 20:404–410. [PubMed: 19478034]
- Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*. 2002; 17(2):825–841. [PubMed: 12377157]
- Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. *Med Image Anal*. 2001; 5:143–156. [PubMed: 11516708]
- Kelley WM, Macrae CN, Wyland CL, Caglar S, Inati S, Heatherton TF. Finding the self? An event-related fMRI study. *J Cogn Neurosci*. 2002; 14(5):785–794. [PubMed: 12167262]
- Kern RS, Green MF, Fiske AP, Kee KS, Lee J, Sergi MJ, et al. Theory of mind deficits for processing counterfactual information in persons with chronic schizophrenia. *Psychological Medicine*. 2009; 39(4):645–654. [PubMed: 18694537]
- Kerr SL, Neale JM. Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance? *J Abnorm Psychol*. 1993; 102(2):312–318. [PubMed: 8315144]
- Langdon R, Coltheart M, Ward PB, Catts SV. Mentalising, executive planning and disengagement in schizophrenia. *Cognitive Neuropsychiatry*. 2001; 6:81–108.
- Leslie AM. Pretense and Representation - the Origins of Theory of Mind. *Psychological Review*. 1987; 94(4):412–426.
- Mazza M, De Risio A, Surian L, Roncone R, Casacchia M. Selective impairments of theory of mind in people with schizophrenia. *Schizophr Res*. 2001; 47(2–3):299–308. [PubMed: 11278148]
- Milev P, Ho BC, Arndt S, Andreasen NC. Predictive values of neurocognition and negative symptoms on functional outcome in schizophrenia: a longitudinal first-episode study with 7-year follow-up. *Am J Psychiatry*. 2005; 162(3):495–506. [PubMed: 15741466]
- Mitchell JP, Macrae CN, Banaji MR. Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*. 2006; 50:655–663. [PubMed: 16701214]
- Mitchley NJ, Barber J, Gray YM, Brooks N, Livingston MG. Comprehension of irony in schizophrenia. *Cognitive Neuropsychiatry*. 1998; 3:127–137.
- Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. [Research Support, Non-U.S. Gov't]. *Lancet*. 1997; 349(9063):1436–1442. [PubMed: 9164317]
- Penn DL, Corrigan PW, Bentall RP, Racenstein JM, Newman L. Social cognition in schizophrenia. *Psychol Bull*. 1997; 121(1):114–132. [PubMed: 9000894]
- Penn DL, Ritchie M, Francis J, Combs D, Martin J. Social perception in schizophrenia: the role of context. *Psychiatry Res*. 2002; 109(2):149–159. [PubMed: 11927140]
- Pickup GJ, Frith CD. Theory of mind impairments in schizophrenia: symptomatology, severity and specificity. *Psychol Med*. 2001; 31(2):207–220. [PubMed: 11232909]
- Pijnenborg GH, Withaar FK, Evans JJ, van den Bosch RJ, Timmerman ME, Brouwer WH. The predictive value of measures of social cognition for community functioning in schizophrenia: implications for neuropsychological assessment. [Research Support, Non-U.S. Gov't]. *Journal of the International Neuropsychological Society : JINS*. 2009; 15(2):239–247. [PubMed: 19203437]
- Premack D, Woodruff G. Chimpanzee problem-solving: a test for comprehension. *Science*. 1978; 202(4367):532–535. [PubMed: 705342]

- Roncone R, Falloon IR, Mazza M, De Risio A, Pollice R, Necozone S, et al. Is theory of mind in schizophrenia more strongly associated with clinical and social functioning than with neurocognitive deficits? *Psychopathology*. 2002; 35(5):280–288. [PubMed: 12457019]
- Russell TA, Rubia K, Bullmore ET, Soni W, Suckling J, Brammer MJ, et al. Exploring the social brain in schizophrenia: left prefrontal underactivation during mental state attribution. *Am J Psychiatry*. 2000; 157(12):2040–2042. [PubMed: 11097974]
- Samson D, Apperly IA, Chiavarino C, Humphreys GW. Left temporoparietal junction is necessary for representing someone else's belief. *Nat Neurosci*. 2004; 7(5):499–500. [PubMed: 15077111]
- Sarfati Y, Hardy-Bayle MC, Besche C, Widlocher D. Attribution of intentions to others in people with schizophrenia: a non-verbal exploration with comic strips. *Schizophr Res*. 1997; 25(3):199–209. [PubMed: 9264175]
- Saxe R, Kanwisher N. People thinking about thinking people. The role of the temporo-parietal junction in “theory of mind”. *Neuroimage*. 2003; 19(4):1835–1842. [PubMed: 12948738]
- Saxe R, Moran JM, Scholz J, Gabrieli J. Overlapping and non-overlapping brain regions of theory of mind and self reflection in individual subjects. *Soc Cog Affect Neurosci*. 2006; 1:229–234.
- Saxe R, Powell LJ. It's the thought that counts: specific brain regions for one component of theory of mind. *Psychological Science*. 2006; 17(8):692–699. [PubMed: 16913952]
- Saxe R, Wexler A. Making sense of another mind: the role of the right temporo-parietal junction. *Neuropsychologia*. 2005; 43(10)
- Schenkel LS, Spaulding WD, Silverstein SM. Poor premorbid social functioning and theory of mind deficit in schizophrenia: evidence of reduced context processing? *J Psychiatr Res*. 2005; 39(5): 499–508. [PubMed: 15992559]
- Sergi MJ, Green MF. Social perception and early visual processing in schizophrenia. *Schizophr Res*. 2003; 59(2–3):233–241. [PubMed: 12414080]
- Siegal M, Varley R. Neural systems involved in “theory of mind”. *Nat Rev Neurosci*. 2002; 3(6):463–471. [PubMed: 12042881]
- Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp*. 2002; 17(3):143–155. [PubMed: 12391568]
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, et al. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*. 2004; 23(Suppl 1):S208–219. [PubMed: 15501092]
- van Overwalle F. Social cognition and the brain: a meta-analysis. *Human Brain Mapping*. 2009; 30:829–858. [PubMed: 18381770]
- Walter H, Ciaramidaro A, Adenzato M, Vasic N, Ardito RB, Erk S, et al. Dysfunction of the social brain in schizophrenia is modulated by intention type: an fMRI study. *Soc Cog Affect Neurosci*. 2009; 4(2):166–176.
- WHO. The global burden of disease: 2004 update. Geneva: 2008.
- Woolrich MW, Behrens TE, Beckmann CF, Jenkinson M, Smith SM. Multilevel linear modeling for fMRI group analysis using Bayesian inference. *Neuroimage*. 2004; 21:1732–1747. [PubMed: 15050594]
- Woolrich MW, Ripley BD, Brady M, Smith SM. Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage*. 2001; 14:1370–1386. [PubMed: 11707093]
- Worsley, KJ. Statistical analysis of activation images. In: Jezzard, P.; Matthews, PM.; Smith, SM., editors. *Functional MRI: An Introduction to Methods*. USA: Oxford University Press; 2001.
- Zaitchik D. When representations conflict with reality: the preschooler's problem with false beliefs and “false” photographs. *Cognition*. 1990; 35(1):41–68. [PubMed: 2340712]

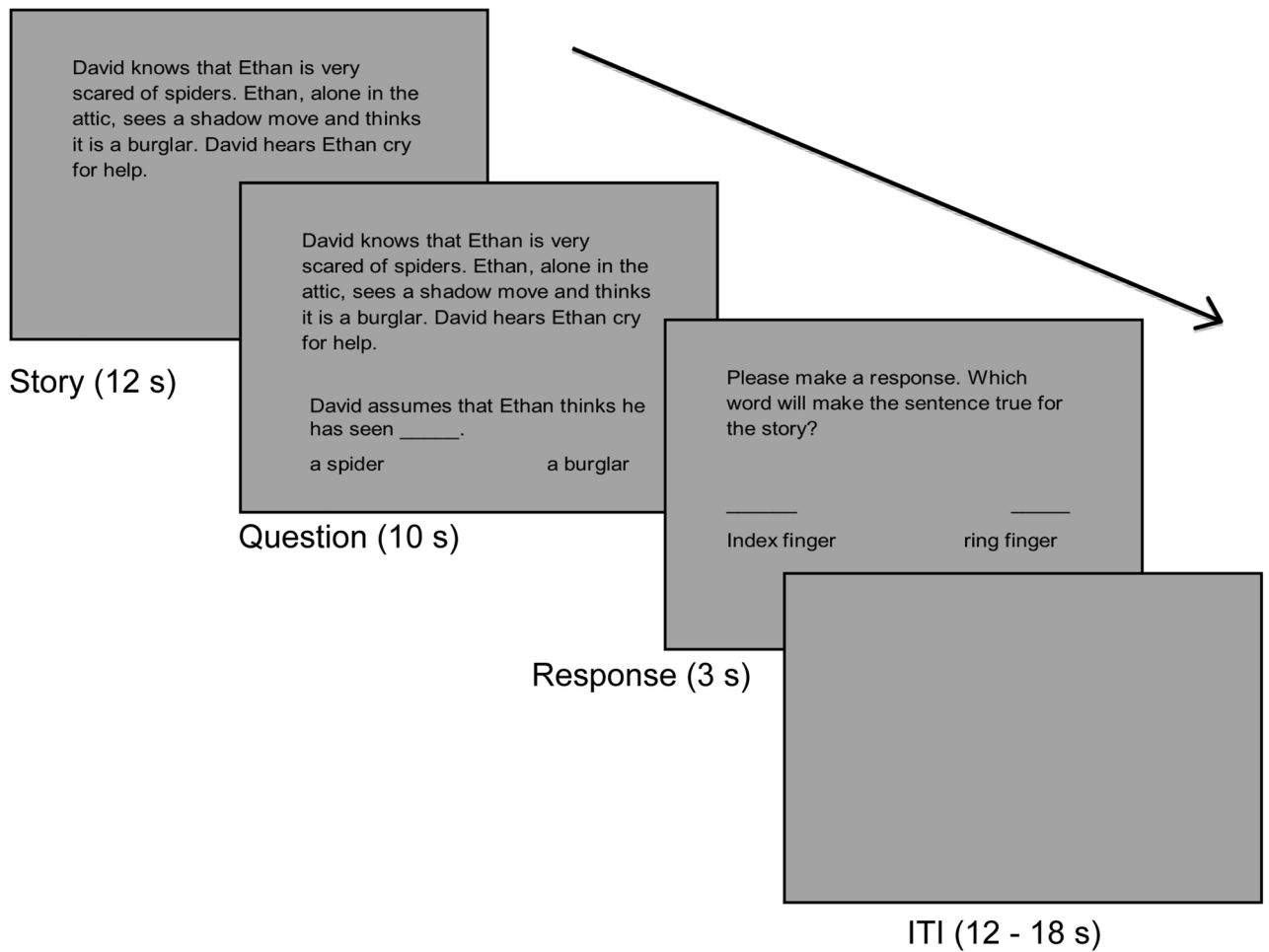
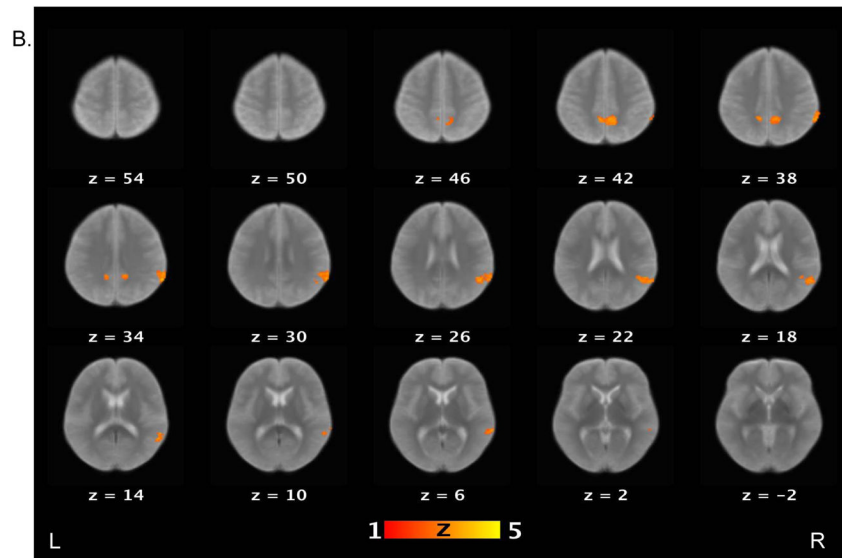
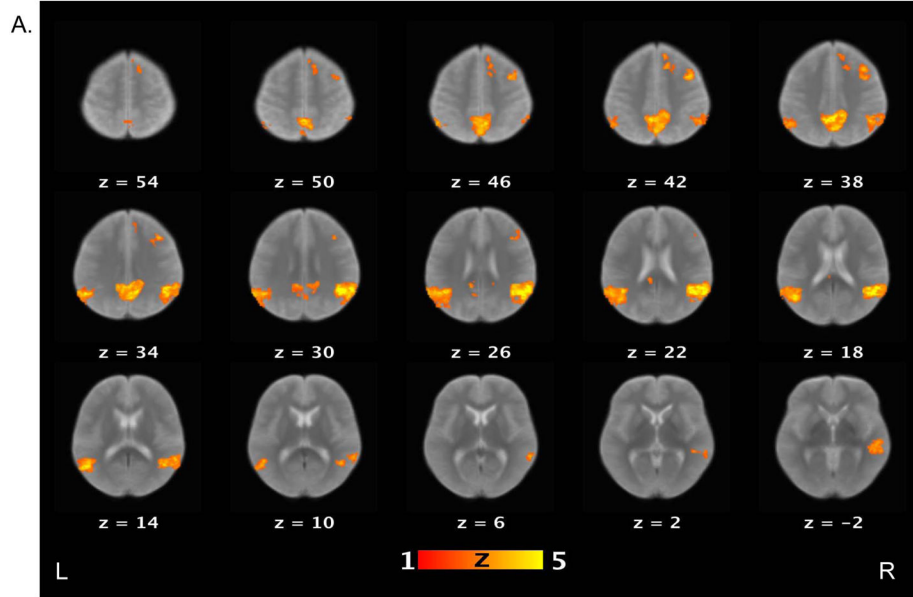


Figure 1.

A schematic diagram of the Belief Attribution Task showing the temporal sequence of a single event. At the beginning of each trial, a vignette was presented for 12 seconds. A fill-in-the-blank question was presented for 10 seconds while the vignette was still visible. After the vignette and question disappeared, a response probe was presented for 3 seconds. The inter-trial interval (ITI) was jittered between 12 and 18 seconds.



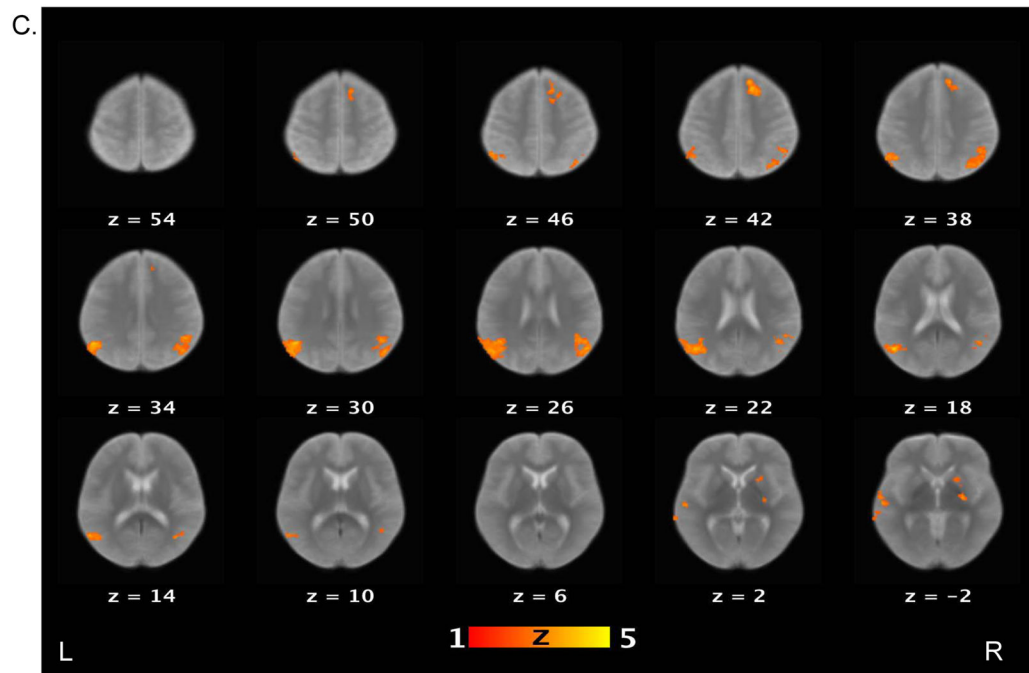
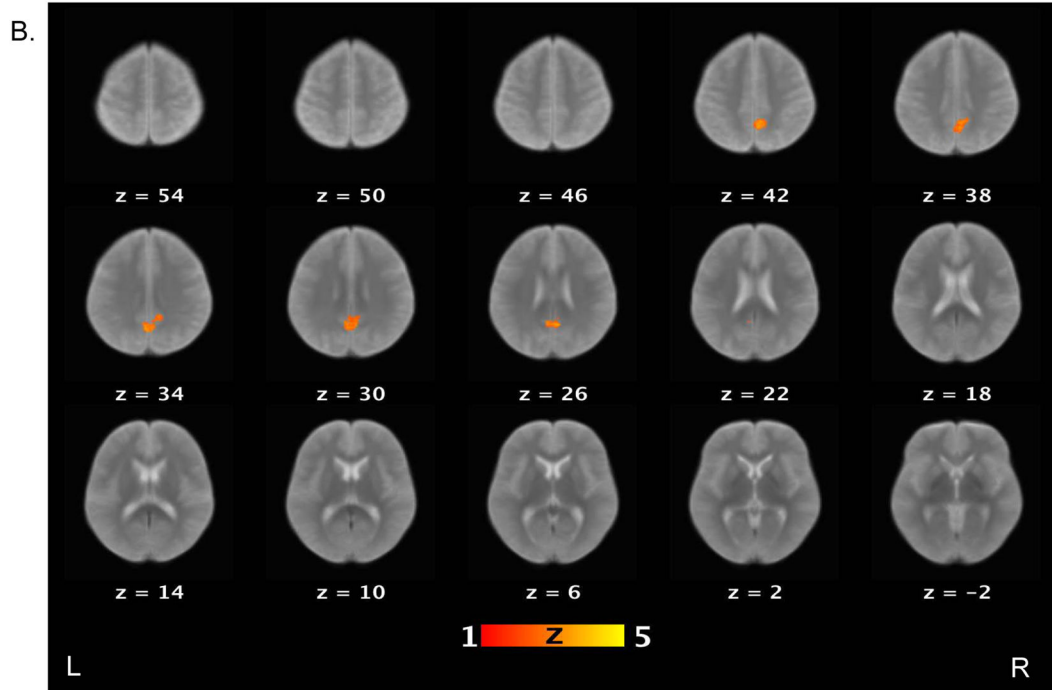
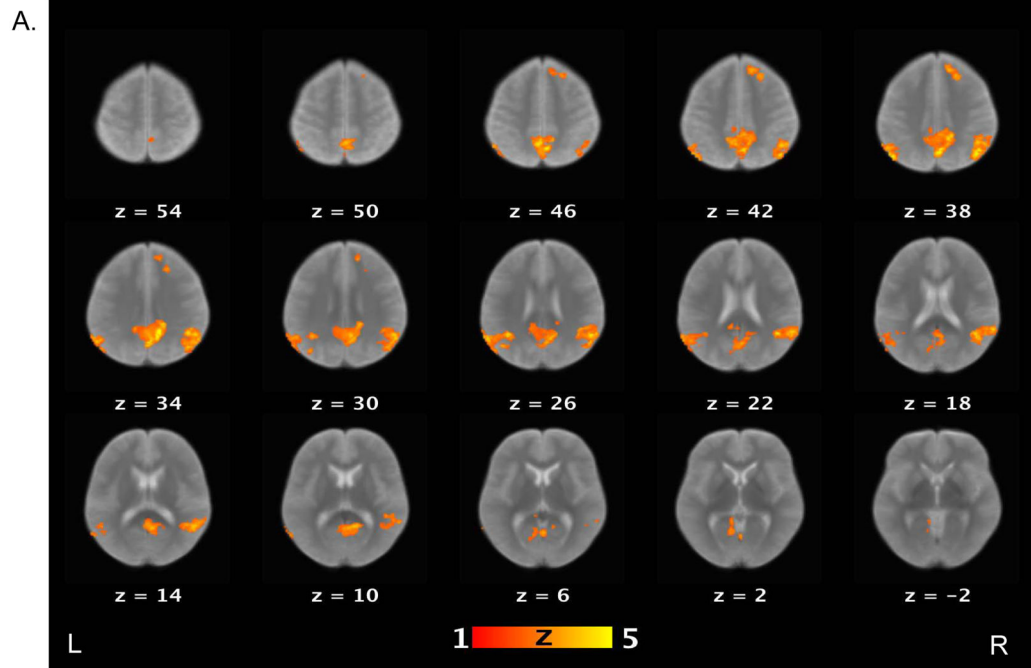


Figure 2. Brain activations for the contrast of false belief versus simple reading conditions. Sections of brain templates with overlaid group analysis results of significant increase in signal intensity during false belief versus simple reading conditions in (A) controls, (B) schizophrenia patients, and (C) controls versus schizophrenia patients.



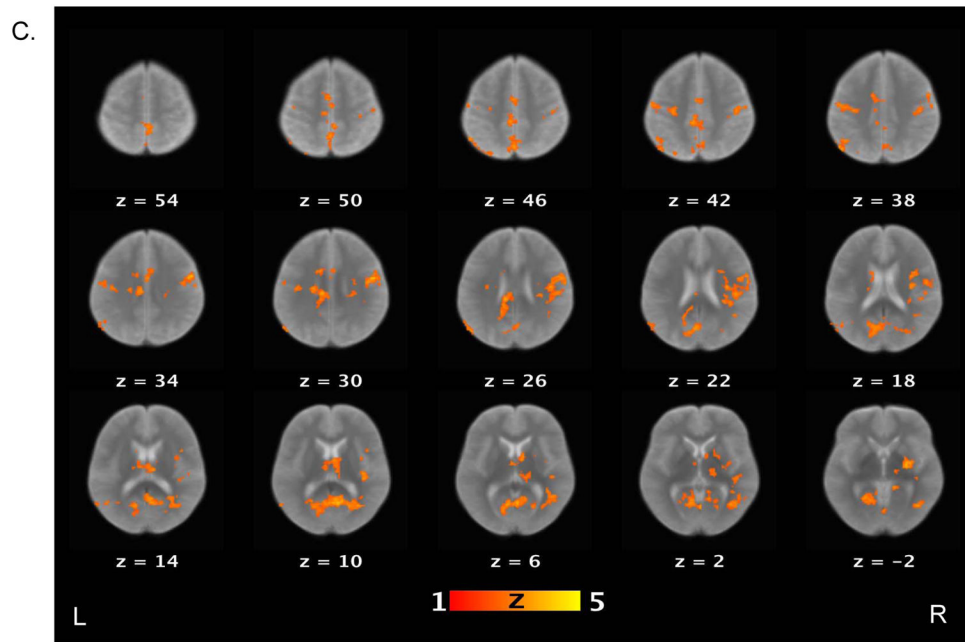


Figure 3. Brain activations for the contrast of false belief versus false photograph conditions. Sections of brain templates with overlaid group analysis results of significant increase in signal intensity during false belief versus false photograph conditions in (A) controls, (B) schizophrenia patients, and (C) controls versus schizophrenia patients.

Table 1

Examples of vignettes

Condition	Vignette	Question
False Belief	David knows that Ethan is very scared of spiders. Ethan, alone in the attic, sees a shadow move and thinks it is a burglar. David hears Ethan cry for help.	David assumes that Ethan thinks he has seen _____. a spider a burglar
False Photograph	Amy made a drawing of a tree house three years ago. That was before the storm. We build a new tree house last summer, but we painted it red instead of blue.	The tree house in Amy's drawing is _____. red blue
Simple Reading	A lemon tree can grow up to 20 ft, but they are usually smaller. The branches are thorny, and the leaves are green and shiny. Flowers are white on the outside with a violet streaked interior.	The flower of a lemon tree is _____ on the outside. white violet

Table 2

Demographic information and Behavioral performance

	Schizophrenia patients (N=12)	Healthy controls (N=13)
Age	38.3 (10.7)	42.5 (7.7)
Gender (female / male)	2 / 10	3 / 10
Personal Education	12.5 (2.3)	14.2 (1.3)
Parental Education	13.1 (2.9)	14.3 (2.9)
Belief Attribution Task *		
False belief	8.5 (1.3)	10.5 (1.3)
False photograph	7.6 (1.8)	9.4 (1.1)
Simple reading	9.9 (1.5)	11.2 (0.7)

† Values are given as mean (standard deviation).

* accuracy = number of correct responses out of 12

Table 3

Brain areas that showed significant activation above threshold for the contrasts of interests

Areas	Hemisphere		x	y	z	z stats
<u>False Belief vs. Simple Reading</u>						
Controls	angular gyrus	R	42	-54	24	7.08
	angular gyrus	L	-40	-54	24	5.34
	medial prefrontal cortex	R	14	26	42	3.58
	middle frontal gyrus	R	40	10	42	4.37
	middle temporal gyrus	R	60	-26	-8	5.14
	middle temporal gyrus	L	-52	-60	14	5.33
	middle temporal gyrus	L	-64	-18	-8	4.26
	precentral gyrus	R	38	22	34	3.89
	precuneus	L	-8	-54	40	6.72
	precuneus	R	4	-56	34	5.58
Patients	superior frontal gyrus	R	10	40	42	3.37
	superior temporal gyrus	L	-44	-54	24	5.22
	temporo-parietal junction	R	50	-56	26	7.75
	temporo-parietal junction	L	-52	-58	34	5.17
	precuneus	R	12	-56	44	3.72
	precuneus	L	-6	-50	42	3.55
	superior temporal gyrus	R	66	-42	6	3.74
	supramarginal gyrus	R	62	-54	24	3.53
	temporo-parietal junction	R	52	-54	24	4.3
	temporo-parietal junction	R	54	-54	18	3.71
Controls > Patients	amygdala	R	24	-2	-12	3.27
	globus pallidus	R	28	-14	-2	3.63
	medial prefrontal cortex	R	14	36	40	3.8
	middle temporal gyrus	L	-48	-66	18	4.45
	middle temporal gyrus	L	-62	-32	-12	4.23
	putamen	R	22	12	-10	3.28
	superior frontal gyrus	R	20	24	42	3.28
	superior lateral occipital cortex	L	-54	-68	32	4.16

Areas	Hemisphere	x	y	z	z stats
superior lateral occipital cortex	R	52	-64	26	3.99
superior temporal gyrus	L	-52	-58	34	4.37
supramarginal gyrus	R	52	-54	36	3.92
temporo-parietal junction	L	-50	-58	30	4.52
temporo-parietal junction	R	48	-54	30	3.58
<u>False Photograph vs. Simple Reading</u>					
Controls					
middle temporal gyrus	L	-58	-58	16	3.52
superior temporal gyrus	L	-56	-48	18	3.99
supramarginal gyrus	R	52	-50	30	4.12
temporo-parietal junction	R	56	-54	26	4.17
temporo-parietal junction	L	-56	-56	26	4.6
middle frontal gyrus	R	36	42	16	3.31
superior frontal gyrus	R	30	56	18	3.85
<u>False Belief vs. False Photograph</u>					
Controls					
angular gyrus	R	44	-66	38	4.94
angular gyrus	L	-50	-70	40	4.83
angular gyrus	L	-48	-70	44	3.81
cuneus	R	4	-64	38	5.12
medial prefrontal cortex	R	22	28	36	4.01
middle temporal gyrus	L	-44	-68	26	3.9
middle temporal gyrus	R	62	-8	-14	4.39
posterior cingulate gyrus	R	18	-42	32	4.75
precuneus	R	12	-52	34	5.35
superior frontal gyrus	R	14	38	40	4.09
superior lateral occipital gyrus	R	54	-62	28	4.65
superior temporal gyrus	R	42	-52	22	4.9
superior temporal gyrus	L	-38	-54	26	4.51
superior temporal gyrus	R	46	12	-30	3.54
temporo-parietal junction	R	56	-54	24	4.87
temporo-parietal junction	L	-60	-58	26	4.28
cingulate gyrus	L	-4	-56	30	3.37
Patients					

Areas	Hemisphere	x	y	z	z stats
posterior cingulate cortex	R	6	-56	26	3.6
precuneus	L	-2	-62	34	3.83
precuneus	R	6	-50	40	3.75
angular gyrus	L	-48	-72	38	3.48
cingulate gyrus	L	-8	-22	32	3.74
cingulate gyrus	R	4	8	32	3.32
culmen	R	26	-46	-16	3.42
fusiform gyrus	R	32	-46	-16	3.17
globus pallidus	R	28	-12	-2	3.84
inferior parietal lobule	L	-52	-58	46	3.18
medial prefrontal cortex	L	-2	0	50	3.05
middle temporal gyrus	L	-48	-66	16	3.17
paracentral lobule		0	-12	48	3.04
parahippocampus	R	32	-10	-20	3.98
postcentral gyrus	L	-38	-16	32	3.34
postcentral gyrus	R	50	-2	30	4.26
posterior cingulate cortex	R	8	-58	8	4.3
precentral gyrus	L	-50	-12	36	3.12
precuneus		0	-70	46	3.38
precuneus	L	-28	-78	48	3.24
superior lateral occipital gyrus	L	-52	-74	26	3.36
thalamus	R	4	-12	12	3.81
thalamus	L	-8	-12	12	3.45

Controls > Patients