Neurobiological correlates of theory of mind in psychosis proneness
Modinos, Gemma; Renken, Remco; Shamay-Tsoory, Simone G.; Ormel, Johan; Aleman, Andre
Published in:
Neuropsychologia

DOI:
10.1016/j.neuropsychologia.2010.09.030

IMPORTANT NOTE: You are advised to consult the publisher’s version (publisher’s PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher’s PDF, also known as Version of record

Publication date:
2010

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Neurobiological correlates of theory of mind in psychosis proneness

Gemma Modinos, Remco Renken, Simone G. Shamay-Tsoory, Johan Ormel, André Aleman

1. Introduction

Theory of mind (ToM) refers to the ability to infer one’s own and other persons’ mental states. ToM abilities are compromised in schizophrenia, in association with dysfunctional activity in predominantly prefrontal brain regions. Prior behavioral studies have also suggested ToM deficits in healthy individuals with psychosis proneness (PP), although no study to date had investigated the associated neural mechanisms in such a sample. Here we used functional magnetic resonance imaging (fMRI) to compare brain activation of subjects with high versus low scores on positive-dimension PP and a ToM task. The ToM task involved first and second order attribution of cognitive and affective mental states to a cartoon character based on verbal and eye-gaze cues. No between-group differences were found on behavioral performance. fMRI analyses revealed a group interaction in anterior prefrontal cortex (BA 10), with the high PP group showing significantly more activity thereof, relative to the low PP, during second order mentalizing. Further between-group differences were observed in dorso-medial and lateral prefrontal regions (BA 46/9), with the high PP group also showing greater activation during second order mentalizing. These results suggest that subjects with positive-dimension PP require more activation of prefrontal areas to adequately mentalize. Differences in the neural mechanisms underlying ToM might be associated with vulnerability to psychosis.

© 2010 Elsevier Ltd. All rights reserved.
although the accumulated findings of ToM deficits in remitted and non-remitted patients suggest that there is indeed a trait-related ToM impairment in schizophrenia, more careful research is needed to investigate this ability in symptom-free patients and people at risk for developing psychosis.

Schizotypy describes a continuum of personality characteristics and experiences related to psychosis in the general population (Claridge et al., 1996). Schizotypal traits can be psychometrically identified in healthy people (Claridge, 1997; Lenzenweger, 1994; Stefanis et al., 2002). There is evidence to suggest that schizotypal traits fall into a factor organization similar to that in schizophrenia, consisting of positive (e.g., magical ideation, perceptual aberration), negative (e.g., physical anhedonia, social anhedonia), and disorganized (e.g., disorganized speech and behavior) symptom dimensions (Claridge et al., 1996; Kerns, 2006; Liddle, 1987). In fact, several studies that have investigated unaffected relatives of patients and samples from the general population provide compelling evidence for continuity between subclinical (e.g., schizotypal traits) and clinical (e.g., schizophrenia) forms of psychosis. For instance, Vollema, Sitksnoorn, Appels, & Kahn (2002) studied relatives of patients with schizophrenia and reported that the risk percentage for the development of schizophrenia was reflected in the score on the positive-dimensional scale of a schizotypal personality questionnaire, which suggests that positive schizotypy reflects the biological–genetic vulnerability to schizophrenia. A large general population twin study investigating 3685 individuals, including 1438 complete twin pairs, found evidence for familial resemblance and a genetic effect for both positive and negative schizotypy dimensions as measured psychometrically (Hay et al., 2001), which represented a replication of the results found with clinical samples. Furthermore, Fanous, Gardner, Walsh, and Kendler (2001) reported that positive symptoms in probands with non-affective psychosis were predictive of positive schizotypy in the relatives, while negative symptoms in the probands were predictive of negative schizotypy, which further suggests continuity of psychotic symptoms. Thus, there is evidence for familial, and possibly genetic, homotypy of these psychosis dimensions (Myin-Germeys, Krabbendam, & Van Os, 2003).

Schizotypal traits are therefore thought to constitute a range of enduring, biologically determined, personality and cognitive traits that predispose to schizophrenia (Chapman, Chapman, Kwapil, Eckblad, & Zinser, 1994; Lenzenweger, 2006). The detection of schizotypal traits in healthy subjects is used as an indicator of psychosis proneness (PP) (Meyer & Hautzinger, 2002), which is conceptualized as a subclinical manifestation of the same underlying biological factors of schizophrenia-spectrum disorders (Johns & van Os, 2001; Van Os, Linscott, Myin-Germeys, Delespaual, & Krabbendam, 2009). Indeed, several prospective studies have shown that about 10% of subjects psychologically identified as psychosis-prone will go on to develop a schizophrenia-spectrum disorder (Chapman et al., 1994; Hanssen, Bak, Bijl, Vollebergh, & Van Os, 2005; Meehl, 1990; see Van Os et al., 2009 for review). A high score on a questionnaire measuring schizotypal personality traits can therefore be conceptualized as a phenotypic marker of risk for schizophrenia (Chapman et al., 1994; Squires-Wheeler, Skodol, & Erlenmeyer-Kimling, 1991). In particular, scales measuring positive schizotypy demonstrate characteristics of vulnerability indicators to schizophrenia and schizophrenia-spectrum disorders (Horan, Blanchard, Clark, & Green, 2008).

Research in PP has revealed impairments on measures of emotional, social and cognitive functioning parallel to those of schizophrenia patients (Henry et al., 2009; Horan, Reise, Subotnik, Ventura, & Nuechterlein, 2008; Mohanty et al., 2008; Mohanty et al., 2005; Van’T Wout, Aleman, Kessels, Laroi, & Kahn, 2004), as well as in brain function and structure (Modinos, Mechelli, et al., 2010). Thus far there is evidence to suggest a mentalizing disturbance in people with PP (see Sprong et al., 2007), in line with the notion that it may be an underlying marker of vulnerability. Such evidence has been commonly provided by studies which have not differentiated between symptom dimensions in PP, with the exception of Pickup (2006), who reported significant ToM deficits in association with the positive rather than with the negative dimension. Moreover, subjects with positive-dimension PP (experiencing e.g., unusual beliefs and aberrant perceptions) show elevated social anxiety and poorer social functioning (Brown, Silvia, Myin-Germeys, Lewandowski, & Kwapil, 2008; Kwapil, Barrantes-Vidal, & Silvia, 2008), which is thought to be related to ToM deficits. Patients with schizophrenia show functional (see Brunet-Gouet & Decety, 2006 for review) and structural brain abnormalities in ToM-relevant regions (Benedetti et al., 2009; Herold et al., 2009; Yamada et al., 2007). Interestingly, a previous fMRI study in individuals at genetic risk for schizophrenia reported abnormal brain activation in prefrontal regions relevant to ToM processing (Marjoram, Job, et al., 2006). In fact, it has been recently suggested that functional and structural abnormalities within brain regions dedicated to self and other-referential processing may be implicated early in the pathophysiology of the disorder (Nelson et al., 2009). To date, however, no study has examined brain activation during ToM in psychosis-prone individuals. Research on such a sample has several strengths, as it allows for the study of mechanisms relevant to psychotic experiences without the confounding factors of medication, illness duration, institutionalization or other consequences of the clinical disorder.

Here we used functional magnetic resonance imaging (fMRI) to examine brain function associated with ToM in a group of individuals with high positive-dimension PP, comparing them with a group of subjects with low positive-dimension PP. In light of recent evidence that tasks involving inference regarding cognitive mental states and tasks involving inference regarding affective mental states are differentially impaired in individuals with schizophrenia (Shamy-Tsoory, Aharon-Perez, & Levkovizt, 2007), we adapted a task that had previously allowed for the study of these components in schizophrenia (Shamy-Tsoory, Shur, et al., 2007). We tested the hypothesis that high positive-dimension PP individuals would show differences in activation, relative to low positive-dimension PP individuals, in prefrontal regions involved in ToM during the correct attribution of mental states, consistent with the one available fMRI study in high-risk relatives of patients with schizophrenia (Marjoram, Job, et al., 2006).

2. Materials and methods

2.1. Participants

Six hundred undergraduate students were screened with the positive subscale of the Community Assessment of Psychic Experiences questionnaire (CAPE; Stefanis et al., 2002). They all gave written informed consent to complete the CAPE. According to their CAPE scores, 36 subjects were ultimately recruited for the actual fMRI experiment. Eighteen right-handed individuals with a high score on the CAPE positive dimension (above the 75th percentile, as recommended in Konings, Bak, Hansen, van Os, & Krabbendam, 2006) were assigned to the “high PP” group (10 men, mean age 19.8 ± 1.9 years, range 18–24, mean CAPE positive-dimension score 1.74 ± 0.13), and 18 right-handed individuals scoring below the 25th percentile of the distribution were included in the low psychosis-prone group (“low PP”: 10 men, mean age 21 ± 2.8 years, range 18–27, mean CAPE positive-dimension score 1.12 ± 0.04), thus, groups were matched for age, sex, handedness, and level of education. These subjects were screened for exclusion criteria using a self-report checklist for healthy subjects, comprising the following points: (1) no personal history of neurological or psychiatric illness; (2) no family history of psychotic or neurological illness in first-degree relatives; (3) no use of illicit substances; and (4) no changes in overall level of functioning, including academic performance over the past 6 months. All 36 participants gave written informed consent for participating in the fMRI experiment after a detailed explanation of the experimental protocol, approved by the Medical Ethical Committee of the University Medical Center Groningen. The procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation and with the Helsinki Declaration of 1975, as revised in 1983.
Fig. 1. Examples of task trials for each condition and order. Text translates to, from left to right and top to bottom: “Jan thinks of...” (first order cognitive); “Jan thinks of the flower that... wants” (second order cognitive); “Jan loves...” (first order affective); “Jan loves the fruit that... loves” (second order affective); “Jan is near...” (first order physical); “Jan has the same toy as...” (second order physical).

2.2. Psychosis proneness questionnaire

The CAPE was used to measure PP. This instrument was chosen based on the following characteristics: (1) good validity and reliability for the assessment of schizotypal features in the general population (Hanssen et al., 2005), (2) good concurrent validity with interview-based measures (Konings et al., 2006), (3) developed and standardized on a Dutch population. A detailed description of its psychometric properties and administration procedure can be found elsewhere (Stefanis et al., 2002). In brief, this is a 42-item self-report questionnaire measuring life-time frequency of attenuated psychotic symptoms, on a 4-point scale of “never”, “sometimes”, “often” and “nearly always”. Konings et al. (2006) reported high effect sizes for the internal stability of the CAPE (0.6–0.8), indicating that self-reported dimensions of psychosis proneness at baseline were strongly associated with the same dimensions at follow-up. Thus, the time lag between completion of the CAPE and time of scanning would not affect the findings.

Originally, the CAPE was based on a three-factor structure of positive, negative and depressive factors. For the purpose of the present study, we only used the scores on the positive factor. We followed the notion that some of the positive symptoms of schizophrenia reflect an impairment in the ability to infer the mental states of others (Frith & Corcoran, 1996), and on prior evidence that impaired indices of ToM are predicted by schizotypal experiences analogous to positive symptoms of schizophrenia in healthy subjects (Pickup, 2006), and by attenuated positive symptoms in relatives of patients with schizophrenia (Marjoram, Miller, et al., 2006). Thus, we measured positive factor schizotypal traits (e.g., unusual experiences, odd beliefs), which relate to the positive dimension of the schizotypy concept (Claridge et al., 1996). The positive dimension of the CAPE is significantly correlated with the positive dimension of the Structured Interview for Schizotypy, Revised (SIR-R) ($\beta = 0.52, t = 8.48, p = 0.000$), and the positive dimension of the Brief Psychiatric Rating Scale (BPRS) ($\beta = 0.27, t = 3.54, p = 0.000$) (Konings et al., 2006).

2.3. Task and experimental design

The task was based on that of Shamay-Tsoory, Aharon-Peretz, et al. (2006) and Shamay-Tsoory, Shur, et al. (2006), and involved the ability to judge mental states based on verbal and eye-gaze cues. The task has been validated before and has been shown to be positively correlated with verbal measures of ToM such as false belief stories (Shamay-Tsoory & Aharon-Peretz, 2007). The task was adapted for the fMRI environment as a mixed experimental paradigm, comprising first and second order mental state attribution using color cartoon images. The stimuli were further divided into three conditions: cognitive, affective and physical.

In short, the task consisted of 54 trials, each showing a cartoon outline of a face (named Jan) and four colored pictures of objects belonging to a single category (e.g., fruits, chairs) or faces, one in each corner of the computer screen (see stimuli examples in Fig. 1). The subject was required to indicate the correct answer (the image to which Jan was referring), based on a sentence that appeared at the top of the screen and available cues, such as Jan’s eye gaze, Jan’s facial expression, or the face’s (the one to which Jan is referring) eye gaze and facial expression. Subjects were instructed to select the correct picture using a four-button response box as fast as they could. There were two main conditions: “theory of mind” (40 trials) and “physical” (14 trials) requiring a first (24 trials) or second (30 trials) order inference. The theory of mind condition involved mental inferences, while the physical condition involved mental inferences, while the physical condition required a choice based on a physical attribute of the character (thus serving as control condition, to ensure that the subject understands the task). In the first order physical condition the subject was asked to refer to Jan’s location (Jan is near...), while in the second order condition the subject was asked to point to the same
object Jan and the character have (Jan has the same fruit as _ has). In the theory of mind conditions, the sentences could be, for example, “Jan is thinking of _”, or “Jan loves _”. In the second order condition, the four stimuli consisted of face images and the choice of the correct response required understanding of the interaction between each of these figures and Jan's mental state. For example, “Jan is thinking of the toy that _ wants”, and “Jan loves the toy that _ loves”. Subjects' performance was rated for accuracy and reaction time.

Specifically, the conditions were as follows:

- First order cognitive condition (FOCOG) in which responding to the sentence on top the image required attribution of a cognitive mental state to Jan.
- First order affective condition (FOAFF) in which responding to the sentence on top the image required attribution of an affective mental state to Jan.
- Second order cognitive condition (SOCOG) in which responding to the sentence on top the image required attribution of a cognitive mental state to Jan about another character's mental state.
- Second order affective condition (SOAFF) in which responding to the sentence on top the image required attribution of an affective mental state to Jan about another character's mental state.
- First and second order physical conditions (FOPHY; SOPHY) did not require ToM abilities.

The experiment began with first order (FO) mentalizing blocks. There were four FO blocks in total, each including six trials. All trials displayed a verbal cue on top of the screen, to which the subjects responded by indicating the correct answer (the image Jan was referring to) using a four-button response box. Each trial was shown for 5 s, and trial conditions (FOCOG, FoAFF, FOPHY) were presented in random order within each block. Blocks were intermixed with 20-s rest periods (fixation cross). Next, second order (SO) mentalizing blocks were presented. There were five SO blocks in total, each including six trials. These images were also shown for 5 s, and trial conditions (SOCOG, SOAFF, SOPHY) were, again, randomly presented within each block. Blocks were intermixed with 20-s rest periods (fixation cross). There were 2 rest periods more (fixation cross) at the beginning and at the end of the experiment, representing a total of 9 experimental blocks and 10 fixation periods. Thus, this was a mixed design, with FO and SO conditions presented as blocks, and within each block COG, AFF and PHY trials randomly presented as events. Total duration of the experiment was about 9 min. All participants underwent a training session on the task immediately before fMRI scanning, including 12 trials (see Fig. 2).

Data were preprocessed and analyzed using SPM5 (Wellcome Department of Imaging Neuroscience; http://www.fil.ion.ucl.ac.uk/spm/), running under Matlab 7.4, (The MathWorks Inc.). Standard pre-processing was applied, first with slice time correction, and then realignment to the first volume to correct for inter-scan motion artifacts. After realignment, a mean EPI image was created, which was co-registered with the structural T1 image. Subsequently, images were spatially normalized to the standard stereotactic space defined by the Montreal Neurological Institute (MNI) template. Functional images were then smoothed with a 3D isotropic 8-mm full-width/half-maximum (FWHM) Gaussian kernel. Low-frequency noise was removed by applying a high-pass filter (cut-off of 128 s) to the fMRI time-series at each voxel. Significant hemodynamic changes for each condition were examined using the General Linear Model (Friston et al., 1995). For each condition (FOCOG, FOAFF, FOPHY, SOCOG, SOAFF, SOPHY), the brain response was modeled by convolving each individual trial (5 s) with canonical hemodynamic response function. To identify activity in regions related to ToM conditions, we computed t-contrasts of each ToM condition to the respective Physical condition (e.g., FOCOCG > FOPHY; SOAFF > SOPHY). Following the purpose to examine activity in cerebral areas associated with the correct attribution of mental states, based on the need to ensure that we were as unequivocally as possible capturing activity related to ToM as opposed to e.g. distraction or other unrelated processes that could have led to errors, only correct trials were modeled for the fMRI analysis.

Next, each individual contrast image was entered into a second-level random effects analysis to examine task-related activations across groups, as well as between-group differences (repeated measures ANOVA). We aimed at investigating the effects of Group (low PP, high PP), Condition (COG, AFF, Order (FO, SO), and interactions Group by Condition and Group by Order. A conjunction analysis of COG

Fig. 2. Schematic diagram of the experimental design, which involved four first order (FO) blocks of 30 s each, during which six trials of the cognitive, affective and physical condition were randomly presented. FO blocks were followed by five second order (SO) blocks, also of 30s each, during which six trials of the cognitive, affective and physical condition were randomly presented. Blocks were intermixed with 20-s rest periods consisting of a fixation cross. During fMRI scanning, subjects were asked to read the sentence on top of the screen and to indicate which of the four response options was correct.
and AFF (versus PHY, across groups) was conducted to test for common theory of mind activation relative to the control condition.

Finally, we tested the hypothesis that SO trials would be more cognitively demanding than FO trials (Samson, 2009) by comparing, across groups, neural activation associated with SO relative to that associated with FO [(SOCOG + SOAFF) > (FOCOG + FOAFF)]. For the sake of completeness, we also examined the reversed contrast, FO > SO [(FOCOG + FOAFF) > (SOCOG + SOAFF)].

Statistical maps were thresholded at a level of $p < 0.005$ uncorrected, and voxelwise data were corrected for multiple comparisons by spatial extent of contiguous suprathreshold individual voxels at $p < 0.05$ for a cluster, in line with the one previous fMRI study on ToM in subjects at genetic risk for psychosis (Marjoram, Jothi, et al., 2006). Coordinates are reported in MNI (Montreal Neurological Institute) space. Brain regions were identified with the Anatomical Automatic Labeling Toolbox for SPM (Tzourio-Mazoyer et al., 2002).

3. Results

3.1. Behavioral

A Condition (COG, AFF, PHY) by Group (high PP, low PP) repeated measures ANOVA was conducted on the accuracy scores in each trial for each Order type. In First Order trials, there was no main effect of Condition ($F(2,68) = 2.664, p = 0.092$), Group ($F(1,34) = 1.197, p = 0.282$), or Group by Condition interaction ($F(2,68) < 1, ns$). A repeated measures ANOVA on RT revealed no main effect of Group ($F(1,34) < 1, ns$), or Group by Condition interaction ($F(2,68) = 2.015, p = 0.147$). There was a significant main effect of Condition ($F(2,68) = 9.954, p < 0.001$). Bonferroni post-hoc correction revealed that PHY trials were quicker than COG ($p = 0.019$) and AFF trials ($p = 0.001$).

In Second Order trials, the ANOVA on accuracy scores revealed no main effect of Condition ($F(2,68) = 1.479, p = 0.235$), Group ($F(1,34) < 1, ns$) or Group by Condition interaction ($F(2,68) < 1, ns$). With regard to RT, the ANOVA revealed no main effect of Group ($F(1,34) < 1, ns$), or Group by Condition interaction ($F(2,68) = 108.506, p < 0.001$). Bonferroni post-hoc correction revealed that PHY trials were quicker than COG ($p < 0.001$) and AFF trials ($p < 0.001$). There was also a significant Group by Condition interaction ($F(2,68) = 3.329, p = 0.042$), although post-hoc analysis of between-group differences for each condition did not reach significance (SOCOG, $p = 0.230$; SOAFF, $p = 0.796$; SOPHY, $p = 0.073$); See Table 1 for a complete visualization of behavioral data.

Table 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>Low PP (n = 18)</th>
<th>High PP (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOCOG</td>
<td>1852 ms (222)</td>
<td>1794 ms (314)</td>
</tr>
<tr>
<td>FOAFF</td>
<td>1855 ms (281)</td>
<td>1874 ms (394)</td>
</tr>
<tr>
<td>FOPHY</td>
<td>1789 ms (230)</td>
<td>1674 ms (303)</td>
</tr>
<tr>
<td>SOCOG</td>
<td>2955 ms (356)</td>
<td>3106 ms (385)</td>
</tr>
<tr>
<td>SOAFF</td>
<td>3093 ms (346)</td>
<td>3128 ms (436)</td>
</tr>
<tr>
<td>SOPHY</td>
<td>2328 ms (239)</td>
<td>2147 ms (340)</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOCOG</td>
<td>98.4</td>
<td>98.4</td>
</tr>
<tr>
<td>FOAFF</td>
<td>97.5</td>
<td>98.8</td>
</tr>
<tr>
<td>FOPHY</td>
<td>97.1</td>
<td>97.2</td>
</tr>
<tr>
<td>SOCOG</td>
<td>79.8</td>
<td>81.5</td>
</tr>
<tr>
<td>SOAFF</td>
<td>79.9</td>
<td>80.1</td>
</tr>
<tr>
<td>SOPHY</td>
<td>82.4</td>
<td>85.2</td>
</tr>
</tbody>
</table>

3.2. Functional MRI

3.2.1. Effect of Group (low PP, high PP)

A voxel-wise repeated measures ANOVA revealed no significant activations as effect of group surviving the statistical threshold of $p < 0.05$ cluster-level corrected for multiple comparisons.

3.2.2. Effect of Condition (COG, AFF)

With regard to RT, the ANOVA revealed no main effect of Group ($F(2,68) = 1.197, p = 0.282$), or Group by Condition interaction ($F(2,68) = 0.235, p = 0.796$; SOPHY, $p = 0.073$); See Table 1 for a complete visualization of behavioral data.

Table 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>Side</th>
<th>Area</th>
<th>Z score</th>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive</td>
<td>R</td>
<td>Superior frontal gyrus</td>
<td>4.47</td>
<td>22</td>
<td>6</td>
<td>54</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>Inferior frontal gyrus</td>
<td>4.12</td>
<td>48</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>Insula</td>
<td>3.48</td>
<td>30</td>
<td>22</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>Middle frontal gyrus</td>
<td>3.60</td>
<td>30</td>
<td>44</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>Supramarginal gyrus</td>
<td>3.71</td>
<td>52</td>
<td>32</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Supramarginal gyrus</td>
<td>2.99</td>
<td>60</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Anterior cingulate cortex</td>
<td>3.02</td>
<td>0</td>
<td>36</td>
<td>2</td>
</tr>
<tr>
<td>Affective</td>
<td>R</td>
<td>Cuneus</td>
<td>5.03</td>
<td>16</td>
<td>92</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>Fusiform gyrus</td>
<td>5.01</td>
<td>24</td>
<td>48</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Fusiform gyrus</td>
<td>4.31</td>
<td>24</td>
<td>48</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>Cerebellum</td>
<td>3.50</td>
<td>6</td>
<td>52</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Hippocampus</td>
<td>2.92</td>
<td>26</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td>Conjunction [(Cognitive &gt; Phy) + (Affective &gt; Phy)]</td>
<td>R</td>
<td>Middle frontal gyrus</td>
<td>4.60</td>
<td>54</td>
<td>28</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Middle frontal gyrus</td>
<td>4.44</td>
<td>36</td>
<td>10</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Precentral gyrus</td>
<td>4.13</td>
<td>48</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>Cerebellum</td>
<td>3.87</td>
<td>30</td>
<td>64</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Cerebellum</td>
<td>3.52</td>
<td>8</td>
<td>78</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Superior temporal gyrus</td>
<td>3.41</td>
<td>56</td>
<td>60</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Inferior parietal lobe</td>
<td>2.98</td>
<td>28</td>
<td>58</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>L</td>
<td>Medial frontal gyrus</td>
<td>2.77</td>
<td>4</td>
<td>26</td>
<td>46</td>
</tr>
</tbody>
</table>

Activations were significant at a threshold of $p < 0.05$, cluster-level corrected for multiple comparisons. Coordinates are in Montreal Neurologic Institute (MNI) space. L, left; R, right.
Table 3
Brain regions showing a significant effect of factor order across groups.

<table>
<thead>
<tr>
<th>Order</th>
<th>Side</th>
<th>Area</th>
<th>Z score</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
</table>
| First  | L    | Superior medial gyrus                     | 6.45    | 0   | 62  | 0  
|        | R    | Calcarine gyrus                           | 5.97    | −4  | −90 | 14 
|        | L    | Supplementary motor area                  | 5.25    | 4   | −24 | 56 
|        | R    | Hippocampus/amygdala                      | 5.01    | 26  | −20 | −18 
|        | L    | Posterior cingulate gyrus                 | 3.78    | −6  | −52 | 24 
|        | L    | Temporal pole                             | 3.66    | −48 | 8   | −34 
|        | L    | Precuneus                                 | 3.21    | −20 | −50 | 10 
| Second | R    | Inferior parietal lobe                    | 7.80    | 38  | −46 | 42 
|        | R    | Middle frontal gyrus/inferior frontal gyrus | 7.37   | 48  | 36  | 30 
|        | L    | Cerebellum                                | 6.08    | −12 | −76 | −32 
|        | R    | Cerebellum                                | 3.52    | 16  | −40 | −44 

Activations were significant at a threshold of \( p < 0.05 \), cluster-level corrected for multiple comparisons. Coordinates are in Montreal Neurologic Institute (MNI) space. L, left; R, right.

3.2.3. Effect of Order (FO, SO)

There was a significant effect of FO in the superior medial gyrus, the right supplementary motor area, hippocampus/amygdala, the left calcarine gyrus, fusiform/hippocampus, posterior cingulate cortex, precuneus and temporal pole. SO produced a significant effect in the right inferior parietal lobe, right middle frontal gyrus, inferior frontal gyrus, and the cerebellum bilaterally (see Table 3).

3.2.4. Group by Order interaction

There was a significant Group by Order interaction in the anterior PFC (right superior frontal gyrus, BA10; MNI coordinates xyz, 26 64 6; Z score 3.19). High PP individuals showed greater activation for SO in that region relative to the low PP group, than for FO (Fig. 3). In short, the differences between groups in the anterior medial PFC were highest for SO mentalizing.

Further prefrontal regions of greater activation in high PP for SO mentalizing were identified in the lateral PFC bilaterally (middle frontal gyrus, inferior frontal gyrus, BA9/46), and the right dorsomedial PFC (superior medial gyrus, BA9) (Table 4 and Fig. 4).

Given that RT were shorter for PHY than for the COG and AFF conditions, we sought to examine whether the observed effects at the neural level were reflecting generic task difficulty. To this end, we run the analyses including RT as covariate in the SPM design. The results did not change, with significant effects still observed within the same clusters (albeit of slightly smaller spatial extent), corrected for multiple comparisons.

Finally, we tested the hypothesis that SO trials would be more demanding than FO trials at the neural level (Samson, 2009). To this end, we examined the effects of FO > SO [(FOCOG + FOAFF) > (SOCOG + SOAFF)] and SO > FO [(SOCOG + SOAFF) > (FOCOG + FOAFF)]

Table 4
Brain regions showing significant group differences during second order trials relative to the physical (control) condition.

<table>
<thead>
<tr>
<th>Group contrast and area</th>
<th>Z score</th>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>High PP &gt; low PP</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| R Inferior frontal gyrus| 3.64    | 40  | 36  | 14 
| L Inferior frontal gyrus| 2.94    | −56 | 20  | 30 
| R Superior medial gyrus | 3.36    | 6   | 42  | 48 
| L Middle frontal gyrus  | 3.24    | −42 | 8   | 36 
| R Middle frontal gyrus  | 3.11    | 46  | 14  | 44 
| R Precentral gyrus      | 3.66    | −44 | 0   | 32 

Activations were significant at a threshold of \( p < 0.05 \), cluster-level corrected for multiple comparisons. Coordinates are in Montreal Neurologic Institute (MNI) space. L, left; R, right.
Fig. 4. Images showing activation location and extent in each region in which subjects with high psychosis proneness (PP) exhibited significantly greater activation for second order mentalizing than for first order mentalizing, relative to subjects with low PP. Graphs represent mean BOLD response levels with confidence intervals in each peak voxel. Activations were significant at \( p < 0.05 \), cluster-level corrected for multiple comparisons. IFG: inferior frontal gyrus; LPFC: lateral prefrontal cortex; medPFC: medial prefrontal cortex; MFG: middle frontal gyrus; SMedG: superior medial gyrus.

across groups, using the same factorial design. This revealed, across groups, stronger activation of a large right-sided cluster including voxels within the middle frontal gyrus and the inferior frontal gyrus (BA9/46) for SO > FO (peak MNI coordinates \( xyz = 48, 36, 30; Z \text{ score} = 7.37, \text{ size} = 1500 \)). There were no PFC regions showing increased activation in FO trials relative to SO trials.

4. Discussion

The present study examined the neural circuitry underlying theory of mind in individuals with high positive-dimension psychosis proneness, following the hypothesis that alterations thereof might be associated with vulnerability to psychosis (Bora et al., 2009; Nelson et al., 2009). We report that subjects with high PP showed differences in brain activation in a number of prefrontal cerebral areas during second order mentalizing relative to the comparison group.

Individuals with high positive-dimension PP recruited a region within the anterior PFC (BA 10) to a greater extent for SO trials than for FO trials relative to controls. A recent review of anatomical and brain imaging studies that have looked into the function of BA10 underscored its crucial role in integrating the outcomes of two or more separate cognitive operations in the pursuit of a higher behavioral goal (Ramnani & Owen, 2004). Second order mentalizing was a priori expected to entail higher demands than FO mentalizing across groups (Samson, 2009). Both FO and SO trials involved the processing of language (verbal cue), eye gaze (Jan’s), and mentalizing (about Jan). Nevertheless, SO trials placed higher demands on working memory and attention, as not only Jan’s cues had to be taken into account but also those of the other faces on the screen. Indeed, SO trials relative to FO trials induced increased PFC activation (right middle and inferior frontal gyri) across groups. In this light, the observed between-group differences on the neural level suggest that, although SO trials placed higher demands on PFC than FO trials, high PP subjects required greater effort to reach equivalent performance during SO mentalizing. Given the role ascribed to BA10, our results suggest that high PP subjects needed greater activity to integrate separate cognitive operations in order to correctly mentalize.

This result is in accordance with Marjoram, Job, et al.’s (2006) findings in high-risk relatives of schizophrenia patients. The authors interpreted such increases in PFC activation as evidence of impaired ToM circuitry due to being at enhanced risk for schizophrenia, which could require some form of compensation from additional activity in other brain regions. In addition, increased PFC activation in subjects with positive-dimension PP is in line with Mohanty et al.’s (2005) study on affective interference. The authors reported that individuals with high positive-dimension PP showed significant \( (p < 0.05) \) increases in activation in dorsal and ventral lateral PFC, as well as in limbic regions such as the hippocampus and the amygdala, during maintenance of attentional set in the presence of negative emotional distractors (Emotional Stroop task); with a sample size similar to ours. Such increases were similarly interpreted as indicative of greater effort/compensatory mechanism in order to achieve normal behavioral performance. This is also consistent with prior results from our group of increased PFC activity in light of equal behavioral performance on an emotion regulation task between subjects with high and low scores on positive-dimension PP (Modinos, Ormel, & Aleman, 2010). Of note, after illness onset patients with schizophrenia exhibit abnormal
hemodynamic response in BA10 during mentalizing tasks relative to healthy controls (Brunet-Gouet & Decety, 2006).

Additional increases in activation during SO mentalizing in high positive-dimension PP were identified in dorsomedial and lateral regions of the PFC (BA9/46). Support for the critical role of these PFC areas in ToM processing has been indicated by studies using mentalizing tasks of diverse nature, such as story and cartoon comprehension, as well as the viewing of real-time interaction (Frith & Frith, 2003; Saxe, Carey, & Kanwisher, 2004). Interestingly, positive psychotic symptomatology in subjects at genetic risk for psychosis was reported to be associated with activation differences also located in prefrontal regions, suggesting that psychotic symptoms could arise at least in part from deficits in the neural architecture underlying the mentalizing process required by ToM (Marjoram, Job, et al., 2006). Russell et al. (2000) reported that medicated schizophrenia patients showed less activation in BA9/46 while making errors in the Eyes Task, in which participants were to attribute mental states to photographs of eyes. In light of the lack of significant differences in accuracy scores between our groups, parallel to the behavioral results in Marjoram, Job, et al.’s (2006), the presumed compensatory mechanism appears to be working to a satisfactory extent. An alternative explanation could be that increased activation in high positive-dimension PP reflects a tendency to over-mentalize (Frith, 2004). The experimental paradigm used herein, however, was not set up to explicitly test for this.

Previous overactivation associated with adequate performance on cognitive tasks had been previously reported in patients with schizophrenia (Callicott, Mattay, et al., 2003; Ramsey et al., 2002). Elevated activity in patients was thought to reflect a reduction of the efficiency with which the brain regions that constitute a network communicate with each other (Ramsey et al., 2002). Interestingly, healthy siblings of schizophrenia patients also showed overactivation of PFC regions associated with adequate performance (Callicott, Egan, et al., 2003), which was hypothesized to reflect inefficient prefrontal information processing that increases the risk for schizophrenia. Our findings are seemingly consistent with this notion. Finally, from a structural point of view, regional gray matter changes in regions of the PFC have been detected in psychosis patients (Callicott, Mattay, et al., 2003; Ramsey et al., 2002). Inter-}


ters, positive psychotic symptomatology in subjects at genetic risk for psychosis was reported to be associated with activation differences also located in prefrontal regions, suggesting that psychotic symptoms could arise at least in part from deficits in the neural architecture underlying the mentalizing process required by ToM (Frith & Frith, 1985), which indicates that they measure similar ToM components (Shaywitz & Shaywitz, 2007). With regard to the study sample, this study was confined to subjects with high scores on positive-dimension psychosis proneness, and suggests that they show differential activation of neural systems underlying ToM. However, the present study was not originally designed to test the viability of the trait-marker hypothesis. As recommended by Pousa et al. (2008), longitudinal studies including individuals from both positive and negative dimension PP may help elucidate whether the activation differences occur in association with positive subclinical experiences. Another potential limitation of the present study is that the subjects were recruited from a university sample, thus caution should be used when extrapolating the present findings to the general population. Of note, students functioning at a high level (Meehl, 1962), thus psychosis-prone individuals with high intellectual capacity might cope better with the problems associated with PP (Romme, Honig, Noorthoorn, & Escher, 1992; Van't Wout et al., 2004). This could account at least in part for the absence of behavioral differences. However, the use of undergrad-
bate students is a common strategy that enhances homogeneity of the sample and is consistent with previous studies in psychometrically identified psychosis proneness (e.g., Fernyhough et al., 2008; Langdon & Coltheart, 1999, 2001, 2004; Platek & Gallup, 2002). We did not include a measure of IQ, as one would not expect major intelligence differences between groups selected from a larger sample of healthy subjects at the university level. In addition, ToM deficits in high-positive-dimension PP have been reported to be independent of IQ functioning (Pickup, 2006). On the other hand, IQ is a relevant variable when it comes to theory of mind performance and therefore future studies incorporating a measure of IQ should further illuminate this issue. An advantage of studies in PP is that there is no interference of cognitive impairment with task performance, which may be problematic in patient populations. Finally, the investigation of subjects with negative- or disorganized-dimension PP fell out of the scope of the present experiment. However, further studies into the relation between ToM and other PP dimensions should further expand our findings.

In conclusion, the present study provides evidence of differences in neural activation of prefrontal regions during theory of mind processing between individuals with high and low positive-dimension psychosis proneness. These findings converge with prior evidence for the notion that alterations in ToM circuitry may be associated with vulnerability to psychosis. Thus, the present results suggest that such alterations may reflect pathophysiological mechanisms at play, rather than developing uniquely as a result of illness chronicity or long-term anti-psychotic medication.

Acknowledgments

This research was supported by the European Science Foundation EURYI grant NWO #044035001 to André Almén. We would like to thank Lisette van der Meer for help with the translation of the task into Dutch, and Anita Kuiper for assistance with MRI scanning.

References


