Evidence from paranoid schizophrenia for more than one component of theory of mind

Peter Scherzer1,2*, André Achim1,2, Edith Léveillé1,3, Emilie Boisseau1,4 and Emmanuel Stip5

1 Département de Psychologie, Université du Québec à Montréal, Montréal, QC, Canada, 2 Institut des sciences cognitives, Université du Québec à Montréal, Montréal, QC, Canada, 3 Hôpital du Sacré-Cœur de Montréal et sa Fondation, Montréal, QC, Canada, 4 Hôpital Rivière-des-Prairies, Montréal, QC, Canada, 5 Centre Hospitalier de l’Université de Montréal, Montréal, QC, Canada

We previously reported finding that performance was impaired on four out of five theory of mind (ToM) tests in a group of 21 individuals diagnosed with paranoid schizophrenia (pScz), relative to a non-clinical group of 29 individuals (Scherzer et al., 2012). Only the Reading the Mind in the Eyes Test did not distinguish between groups. A principal components analysis revealed that the results on the ToM battery could be explained by one general ToM factor with the possibility of a latent second factor. As well, the tests were not equally sensitive to the pathology. There was also overmentalization in some ToM tests and under-mentalisation in others. These results led us to postulate that there is more than one component to ToM. We hypothesized that correlations between the different EF measures and ToM tests would differ sufficiently within and between groups to support this hypothesis. We considered the relationship between the performance on eight EF tests and five ToM tests in the same diagnosed and non-clinical individuals as in the first study. The ToM tests shared few EF correlates and each had its own best EF predictor. These findings support the hypothesis of multiple ToM components.

Keywords: schizophrenia, paranoid symptoms, theory of mind, executive functions

INTRODUCTION

Theory of Mind

Theory of mind (ToM) is but one component of social cognition Green et al. (2008). ToM is defined as the ability to attribute, correctly or incorrectly, beliefs, knowledge, feelings or intentions to others, in order to understand and predict their behavior (Perner, 1991; Perner and Lang, 2000; Green et al., 2008). The discovery of this ability coupled with the large inventory of tests used as a measure, allowed researchers to make strides in the understanding of atypical development and specifically in schizophrenia (Scz) (Frith, 1992; Frith and Corcoran, 1996; Langdon and Coltheart, 1999; Champagne et al., 2005; Champagne-Lavau et al., 2006, 2007; Uhlhaas et al., 2006; Martino et al., 2007).

Theory of Mind and Schizophrenia

Research in social cognition in Scz has revealed reliable and large impairments in understanding first and second order false beliefs (Bora et al., 2009a; Bozikas et al., 2011), understanding indirect messages (Corcoran et al., 1995; Greig et al., 2004), inferring affect based on photos of the area around the eyes (Baron-Cohen et al., 2001) identifying irony and faux pas.
(Shamay-Tsoory et al., 2005; Chung et al., 2014), and making inferences concerning real-time social interactions (Bazin et al., 2009; Ouellet et al., 2010; Montag et al., 2011).

**Executive Functions and ToM in Schizophrenia**

Executive functions are considered to be a critical cognitive mediator for ToM (Perner and Lang, 2000). The link between executive deficits and ToM impairment may be explained by difficulty in inhibiting one’s own perspective and distinguishing it from others (Ruby and Decety, 2003). A difficulty in making non-literal interpretations may be due to a difficulty in inhibiting a usual interpretation (Leslie et al., 2004), a lack of flexibility that is reflected in difficulties judging the relative importance of each aspect of a script and attributing the appropriate importance to the pertinent information (Channon and Crawford, 2000).

Significant correlations have been found between a large variety of EF tests and ToM in patients with Scz (see for example Langdon et al., 2001; Bell and Mishara, 2006; Bora et al., 2006a; but see Lysaker et al., 2008 for a contrary opinion and results). Pickup (2008) analyzed 17 studies, eight of which reported a significant correlation between ToM and EF. Although there is a link between the two, he found that the patients were impaired on tests of ToM compared with control subjects, even when EF were factored out. On the other hand, he found that EF shared 65% of the variance on ToM tests in the clinical groups while there was no significant correlation in the non-clinical group (Pickup, 2008).

Bora et al. (2006a) used a battery to probe the relationship between insight into illness, ToM (RMET, first and second order false belief stories) and EF (WCST, Digit Span, verbal fluency, letter-number sequencing). RMET correlated with Digit Span backward and letter-number sequencing, but not WCST. Second order false beliefs correlated with letter-number sequencing, WCST perseveration and categories, but not with Digit Span backward. These results could lead one to consider that there are likely different ToM components as different ToM tests load divergently on EF tests. These discrepancies in results and divergent approaches in the analyses point to inconsistencies in the literature and raise questions about the structure or content of what is being measured.

Montag et al. (2011) identified one ToM component the cognitive/emotional content (see also Shamay-Tsoory et al., 2007). Hynes et al. (2006) identified a differential activation pattern depending on whether the social perspective-taking task was emotional or cognitive, with activation of the medial orbitofrontal lobe distinguishing between the two conditions. Shamay-Tsoory et al. (2006) found that patients with ventromedial prefrontal lesions performed better in the cognitive than in the emotional condition. The differentiation between the two components was further confirmed using ToM tests (Shamay-Tsoory et al., 2006, 2007; See Abu-Akel and Shamay-Tsoory, 2011 for a résumé of the neural circuitry of the components that they identify that include cognitive and emotional components). Salvatore et al. (2008, p. 193, paragraph 1) present an argument, based on data from different populations, in support of a multi-component ToM. However, the clinical evidence extracted from an interview with two patients diagnosed with schizophrenia leaves the debate unresolved.

**Objectives and Hypotheses**

The present study is an attempt to reexamine the link between ToM (faux pas, lies, indirect messages, inferring facial expressions of emotions etc) and EF (cognitive flexibility, deductive reasoning, etc), in a group of patients diagnosed with paranoid schizophrenia (pScz) to determine if different EF measures, are equally good predictors of performance on a battery of ToM tests and if the ToM tests share the same relationship with the EF measures. We predict that although the clinical group will be impaired relative to a non-clinical group on the ToM measures, with the exception of the RMET (Scherzer et al., 2012), that performance on some ToM measures better distinguish between the two groups than others.

We further predict that performance on all the EF tests will be impaired in the pScz group relative to the non-clinical group but the pScz group will perform better on some EF measures than others. Finally, we predict that correlations between the different EF measures and ToM tests, will differ sufficiently within and between groups to support the contention that there is more than one component to ToM.

**MATERIALS AND METHODS**

The ethics committee of the Département de psychologie, Université du Québec à Montréal and the ethics and scientific committee (Comité d’éthique de la recherche) of the Hôpital Louis H. Lafontaine (recently renamed Institut Universitaire en Santé Mentale de Montréal) approved the protocol. Informed written consent was obtained from each subject prior to study entry.

**Participants**

Twenty-one patients diagnosed with pScz and a group of 29 non-clinical individuals were recruited for the study (more details in Scherzer et al., 2012). The subjects were all males, between 18 and 35 years old, either native French speakers or having received all of their schooling in French. Their IQ (VIQ and PIQ, FSIQ) was ≥85. Individuals with Axis I and/or II comorbidity, neurologic problems, head trauma, alcoholism, substance abuse, or dependence, non-corrected visual deficits as determined by the medical records, were excluded from the study. As well, the attending psychiatrists verified that the participants were not under the influence of any recreational drugs or alcohol prior to experimentation.

The diagnosis of pScz was made by the attending psychiatrists and confirmed by ES according to DSM-IV (American Psychiatric Association [APA], 1994) diagnostic criteria. The clinical group was recruited from the outpatient clinic for young adults with psychosis of the Institut Universitaire en Santé Mentale de Montréal. Their medication had to be stable 2 weeks prior to data collection.
The non-clinical group was recruited from the community via posters, word-of-mouth and from talking to groups of individuals. Their socio-demographic profile (age, education, parental education) was comparable to that of the clinical group and their immediate family history (parents, siblings) had to be free of schizophrenia and other psychosis related disorders. They had to be free of Axis I and II comorbidity, neurological problems, head trauma, alcoholism, substance abuse, or dependence, non-corrected visual deficits. This information was obtained during an extensive telephone interview with the potential candidates.

**Measures**

**Clinical Evaluation (Kay et al., 1987)**
The clinical group was evaluated using the Positive and Negative Syndrome Scale (PANSS) during a semi-structured interview. A second psychiatrist independently validated the ratings on the PANSS. The rated >4 on one or more of the following characteristic symptoms on the PANSS: grandiose ideas, delusions, or hallucinations, and persecutory ideation.

**Neuropsychological Measures**

**WAIS-III (Wechsler, 1997)**
An abridged French version of the WAIS-III (Pilgrim et al., 1999) was used to evaluate the intelligence of the participants in order to control for any potential contribution of this variable to the ToM measures. An estimate of VIQ was obtained using the following subtests: Information, Similarities, Digit Span, and Arithmetic. An estimate of PIQ was obtained using the following subtests: Picture Completion, Block Design Substitution.

**Theory of mind tasks**
Theory of mind was measured using five different tests: reading the mind in the eyes test (RMET; Baron-Cohen et al., 2001), Hinting Task (Corcoran et al., 1995; Marjoram et al., 2005), Strange Stories (Happé et al., 1998), and Faux pas (Stone et al., 1998) and Conversations and Insinuations (C and I; Ouellet et al., 2010).

**TABLE 1** | Group comparisons between paranoid schizophrenic (pScz) and non-clinical groups.

| pScz | Non-clinical |
|------|-------------|
| M (SD) | M (SD) |
| Age (yr) | 25.71 (4.44) | 23.07 (3.2) | 2.45* |
| Education, subjects (yrs) | 10.71 (1.55) | 12.03 (1.18) | -3.42** |
| Education, parents (yrs) | 11.94 (2.06) | 12.91 (2.45) | -1.38ns |
| FSIQ | 99.95 (13.58) | 110 (11.04) | 3.11** |
| VIQ | 100.81 (11.84) | 109.07 (14.82) | -2.47** |
| PIQ | 98.24 (12.09) | 111.03 (13.83) | -3.40** |

*Age was positively skewed in pScz (z = 2.45) and negatively skewed in controls (z = -2.45) and acceptably symmetrical in non-clinical controls (z = 1.08). The group difference is still significant by Mann–Whitney U = 134.5, p = 0.001.**

**TABLE 2** | Group comparison on ToM measures scaled for maximum of 100 (adjusted for IQ and education if justified).

| ToM tests | pScz | Non-clinical |
|-----------|------|-------------|
| (n = 21) | (n = 29) |
| M (SD) | M (SD) |
| Hinting task | 72.3 (10.8) | 92.6 (4.0) | 1.48 | 87.66** | 64.6 |
| C and I | 52.3 (14.3) | 73.8 (10.5) | 1.47 | 38.76** | 45.2 |
| Strange stories | 64.6 (21.9) | 94.6 (8.6) | 1.48 | 44.97** | 48.4 |
| Faux Pas | 67.8 (22.8) | 93.2 (10.5) | 1.48 | 48.24** | 50.1 |
| RMET | 56.2 (8.0) | 60.2 (11.0) | 1.48 | 11.18 | 2.4 |

*FSIQ would have been a significant covariate (p = 0.009) but is not included because homogeneity of slope was subsequently rejected (p = 0.009).**

**The back transformed means are 73.7 and 98.1.**
### TABLE 3 | Comparisons between paranoid schizophrenic (pScz) and non-clinical groups’ executive function scores adjusted for IQ and education if justified.

| Test                          | pScz (n=21) M (SD) | Non-clinical (n=29) M (SD) | Df | F Covar | F Group | η² (%) |
|------------------------------|-------------------|---------------------------|----|---------|---------|--------|
| Digit Span (DS) [Scaled Score (SS) forward and backward] | 9.67 (2.13) | 9.00 (2.82) | 47 | 0.904 | 0.83 | 1.7 |
| DS forward (SS)              | 6.72 (1.03) | 6.24 (1.24) | 47 | 0.557 | 1.92 | 3.9 |
| DS backward (SS)             | 4.76 (1.29) | 4.76 (1.55) | 47 | 0.759 | 0.00 | 0.0 |
| Similarities (SS)            | 10.88 (2.33) | 10.19 (3.61) | 48 | 2.72 | 5.4 |
| Trails B-A (sec)             | 41.43 (18.23) | 34.07 (18.07) | 47 | 2.18 | 4.4 |
| Stroop interference (sec)    | 50.73 (10.11) | 50.86 (11.11) | 47 | 1.45 | 3.0 |
| Stroop interference (# errors) | 2.10 (3.03) | 1.35 (2.37) | 47 | 2.56 | 5.2 |
| Stroop flexibility (sec)     | 59.00 (11.32) | 59.62 (16.90) | 46 | 5.75 | 0.36 | 0.1 |
| Stroop flexibility (# errors) | 2.07 (2.62) | 2.82 (2.63) | 46 | 0.89 | 0.28 |
| Fluidity (phonological – # words) | 33.70 (11.08) | 34.55 (11.02) | 46 | 0.70 | 0.1 |
| Fluidity (semantic – # words) | 36.02 (9.12) | 39.41 (7.80) | 45 | 0.77 | 1.58 | 3.4 |
| Hayling control condition (sec) | 42.31 (18.73) | 38.85 (18.81) | 41 | 0.34 | 0.8 |
| Hayling inhibition (# errors) | 8.06 (2.39) | 4.27 (3.17) | 47 | 29.83** | 38.8 |
| Brixton abanal                  | 2.5 (1.86) | 1.79 (1.66) | 42 | 1.68 | 3.8 |
| Tower of London (TOL) (execution time: sec) | 187.95 (79.51) | 192.5 (91.53) | 45 | 5.84 | 0.83 | 1.8 |
| TOL (# errors)                | 0.40 (0.75) | 0.21 (0.42) | 46 | n.a. |
| TOL (planing time: sec)       | 18.10 (19.96) | 50.89 (32.58) | 45.1 | n.a. | t = 4.31** | d = 1.01 |
| TOL (moves)                   | 38.35 (13.65) | 16.68 (11.56) | 46 | 3.256** | 43.4 |
| Zoo map planning time (sec)   | 30.05 (44.63) | 23.79 (26.90) | 46 | 0.28 | 0.6 |
| Zoo Map (# errors)            | 1.05 (0.85) | 0.55 (0.63) | 46 | 5.49* | 10.7 |
| Zoo map (execution time: sec) | 111.74 (62.76) | 40.59 (33.15) | 45 | 5.70 | 14.43** | 24.3 |
| WCST (# categories)           | 3.42 (1.54) | 4.52 (0.92) | 42 | 8.99** | 17.6 |
| WCST (perseveration errors)   | 7.95 (4.6) | 4.96 (4.04) | 42 | 7.06* | 14.4 |
| WCST (non-perseveration errors NP) | 6.84 (5.29) | 4.90 (3.05) | 42 | 4.63* | 9.9 |

*Mean adjusted for FSIQ.

aMean adjusted for years of education.

p < 0.05, **p < 0.01.

bFSIQ covariable excluded for lack of homogeneity of regression slopes (p = 0.04).

cOne control missing. Score root square transformed, no constant added, to reduce skewness from 0.941 (z = 2.13) to 0.356 in controls (0.211 to –0.046 in patients). Back transformed means are 36.35 and 34.00.

dOne patient missing. Because of excess skewness in controls (z = 1.323/0.434 = 3.05), transforming by η²[4570/0.512 = 3.07] made the skewness within on standard error in both groups. Back transformed means are 52.47 and 48.10.

eOne patient missing. Because of excess skewness in controls (z = 1.259/0.434 = 2.90), transforming by η²[4570/0.512 = 3.07] made the skewness within on standard error in both groups. Back transformed means are 2.48 and 1.27.

fOne patient missing.

gOne subject missing in each group.

hOnly available for 16 patients and 27 controls.

iOne patient excluded as an outlier, being at z = –4.7 from the remaining of that group.

jAvailable for 16 patients and 28 controls. Because of excess skewness in controls (z = 1.406/0.441 = 3.19), transforming by η²[4570/0.512 = 3.07] made the skewness within on standard error in both groups. Back transformed means are 2.65 and 1.96.

kAvailable for 20 patients and 28 controls. This variable is positively skewed in both groups: z = 1.572/0.512 = 3.07 for the patients and z = 2.015/0.441 = 4.57 for the controls. The minimum observed was 96. The transformation η²[4570/0.512 = 3.07] brought the skewness index to 0.222 and –0.249, respectively. Tabled raw means do not take the FSIQ covariate into account. Back transformed covaried means are 156.38 and 173.33.

lErrors (0, 1, 2) are distributed 15, 2, 3 for the patients and 22, 6, 0 for the controls. With only two different observed values for the controls, no transformation can change the skewness. Because of this poor distribution, analysis of covariance is not applicable (n.a.). A chi-square test indicates that the two distributions do not differ significantly; χ²(2) = 5.13, p = 0.077.

mTabled means are for 20 patients and 28 controls. But 12 of the patients had the minimum observed score of 10. Since the group variances differ, a t-test for unequal variances is reported. One patient was clearly an outlier for his group, with a score of 97 (quite normal for controls); without this patient, the mean was 13.95 with an SD of 7.5, which makes the outlier some 11 SD above the mean of the remaining patients. Without him, group variance differ even more, the group difference gives t(31.1) = 5.78 and effect size, based on the SD of the controls, is d = 1.13. But with 12 of the remaining 19 patients at the observed minimum, all statistical tests based on normal distributions are of doubtful validity. In the 28 controls, skewness was good at z = 0.180/0.441.

nZoo test not available for two patients. Planning time skewness is not acceptable in the patients (z = 1.678/0.524 = 3.20, with z = 0.814/0.434 = 1.88 in controls). There are several scores of 0 (10 of the 19 patients and 11 of the 29 controls). The transformation η²[4570/0.512 = 3.07] brought skewness to 0.229 and –0.283, respectively. Back transformed means are 4.01 and 5.91.

oSkewness is acceptable in patients (z = 0.118/0.524 but not in controls (z = 1.025/0.434 = 2.38). The observed minimum is 9. The skewness-zoovec = η²[4570/0.512 = 3.07]. Back transformed means are 89.10 and 35.32.

pAvailable for 19 patients and 25 controls. Skewness is negative in both groups and significantly different from zero in controls (z = –0.907/0.524 = –1.73 et z = –1.292/0.464 = –2.78). Scores vary from 0 to 6. Transforming into z-scatsel = 2–η²[4570/0.512 = 3.07] preserves the original polarity without producing negative scores and brings skewness to –0.512 and 0.255, respectively. Back transformed means are 3.63 and 4.67.

qSkewness would be acceptable in patients (z = 0.665/0.524 = 1.27) but is not in controls (z = 1.665/0.464 = 3.59). Range is 1 to 7. Transforming into η²[4570/0.512 = 3.07] brought skewness to 0.540 and –0.396, respectively. Back transformed means are 5.56 and 3.50.
Strange stories (Happé, 1994; Happé et al., 1998). This test consists of eight stories requiring an inference concerning the mental state of a protagonist (Happé et al., 1998) and eight control stories requiring a physical inference. This test was found to distinguish between subjects at high risk for schizophrenia and non-clinical subjects (Chung et al., 2008), as well as between patients with schizophrenia and non-clinical subjects (Langdon and Ward, 2009, 2010).

Faux Pas (Stone et al., 1998; Baron-Cohen et al., 1999). This test consists of 10 stories describing the interaction between two people, one of whom unknowingly makes a comment that is insulting or hurtful, about the other person. It has been found as well to distinguish between schizophrenic patients with negative symptoms and a history of violence (Abu-Akel and Abushu’leh, 2004) as well as between schizophrenic patients with and without insulting or hurtful, about the other person. It has been found that particular variable.

Conversations and insinuations (C and I; Ouellet et al., 2010). This test is composed of four self-contained clips of approximately 2 min duration each, taken from popular French TV programs (see Ouellet et al., 2010 for details). The subject is required to make inferences in order to understand the social interactions, indirect messages, faux pas, white lies, and sarcasm used in the conversation. Each scene is independent of the others and does not require any further information in order to understand the content, nor having viewed previous episodes of the program. See Ouellet et al. (2010) for a more complete description of C and I.

Executive Function Battery

A battery of EF was used to assess speed of processing (Reitan and Wolfson, 1985), cognitive flexibility (Trails B–A, see Lezak et al., 2012) and inhibition (Hayling Sentence Completion Test, Burgess and Shallice, 1997); (Wisconsin Card Sorting Test – Abridged, Kongs, 2000; D-KEFS Stroop, Delis et al., 2001), verbal fluency (D-KEFS Letter and Category Fluency, Delis et al., 2001), Hayling Sentence Completion Test (Burgess and Shallice, 1997), deductive reasoning (Brixton Spatial Anticipation Task, Burgess and Shallice, 1997), problem solving (Tower of London, see Culberston and Zillmer, 2001) and planning (BADS Zoo Map, Wilson et al., 1996).

Procedure

The entire battery required two sessions of approximately 2 h each to administer. The sequence of tests was counterbalanced between subjects. At the end of the testing, each participant received eight dollars in compensation. Participants in the clinical group were tested in the hospital whereas those in the non-clinical group were tested in a university laboratory.

Statistical Analyses

Analyses of covariance (ANCOVA) adjusted for FSIQ and effect size ($\eta^2_p$) were used to compare the performances between groups on each ToM test. Student $t$-tests or analyses of covariance (ANCOVA) adjusted for IQ and education where appropriate, were used to compare the performance of the two groups on the measures derived from the tests of executive functions. Partial Pearson correlations within groups were used to examine the shared variance between EF and ToM measures, after controlling for FSIQ if pertinent.

| ToM test          | pScz group | Non-clinical group |
|-------------------|------------|--------------------|
| Hinting Task      | Digit span backward |                |
| Strange Stories   | TOL errors  |                    |
| Faux Pas          | Zoo Map planning time |          |
| C and I           | Brixton aban.  | Digit Span        |

TOL, Tower of London.

**Table 5** Best executive function predictors of performance on ToM tests.

**Table 4** Correlations within and between pScz and non-clinical groups, between EF and ToM measures, with FSIQ partialled out ($r^p$) when it correlated significantly with that particular variable.

| Group    | EF variable                   | ToM variables                        |
|----------|------------------------------|--------------------------------------|
|          |                              | Hinting Task$^a$          | Strange Stories$^a$       | Faux Pas | C and I$^{ab}$ |
| pScz     | Digit Span backward$^a$       | 0.543*                           | 0.270                   | 0.340    | 0.364        |
|          | Stroop flexibility time      | −0.222                           | −0.617**                | −0.220   | −0.546*      |
|          | Hayling errors               | −0.035                           | −0.239                  | 0.167    | −0.519*      |
|          | Brixton aban.                | −0.452                           | −0.086                  | −0.369   | −0.610*      |
|          | Tower of London errors       | −0.010                           | −0.629**                | −0.020   | 0.327        |
|          | Zoo map planning time        | −0.035                           | 0.141                   | −0.583** | −0.186       |
| Controls | Digit Span SS                | −0.194                           | −0.096                  | 0.076    | 0.441*       |
|          | Digit Span forward           | −0.007                           | −0.115                  | 0.169    | 0.490**      |
|          | Similarities$^b$             | 0.033                            | 0.041                   | 0.411*   | 0.230        |
|          | Traits B–A                   | −0.006                           | 0.001                   | 0.106    | 0.062        |
|          | Tower of London exec. time   | −0.397*                          | 0.085                   | −0.161   | −0.199       |

$^a p < 0.05$, $^b p < 0.01$.

$^a$Correlated with FSIQ in patients.

$^b$Correlated with FSIQ in controls.
RESULTS

Statistical Analyses
All the transformations used are noted in detail in the subtext of the relevant tables, where appropriate.

Between Group Comparisons of Socio-demographic Characteristics
Student t-tests revealed significant group differences for age, education and IQ measures (Table 1).

Comparison between Groups on ToM Tests
Given the group differences in background variables that could influence the ToM or EF variables, the group differences on these variables were tested first by including education and FSIQ as covariates. Only covariates declared significant predictors of the dependent ToM variable at \( p < 0.05 \) were kept and the homogeneity of regression slopes was verified. If homogeneity was rejected (i.e., covariate \( \times \) dependent variable interaction significant at \( p < 0.05 \)), the covariate was excluded and the situation flagged in the tabulated report. The standard deviations reported are the original ones, not reduced by the retained covariates, if any.

Two ToM variables were significantly negatively skewed in controls because of ceiling effects. The skewness index was at \( z = -1.486/0.434 = -3.42 \) for Strange Stories, with 18 of the 29 controls at ceiling. For Faux Pas, skewness was at \( z = -2.958/0.434 = -6.82 \), with 11 at ceiling. For patients, skewness was, respectively, at \( z = -0.713/0.501 = -1.42 \) and \( z = -0.795/0.501 = -1.59 \). Without the participants performing at ceiling, the skewness for Strange Stories reduced to \( z = -0.294/0.661 \), but for Faux Pas it reduced only to \( z = -2.438/0.536 = -4.55 \). The latter value, in association with the negative skewness in the patients, warranted a scale transformation for this variable. The transformation \( L \text{Faux Pas} = 2 - \text{LG10}(60.3 - \text{Faux Pas}) \) brought skewness to \( z = 0.151/0.501 \) for the patients and to \( z = -0.282/0.434 \) all 29 controls and \( z = 0-698/0.536 \) for the 18 controls not at ceiling. The transformed version is used in the statistical analyses, and its group means are reported back transformed in the original scale.

The scores on the ToM tests were scaled for a maximum of 100 to allow for comparisons between tests prior to analyses. The effect size \( (r^2) \) was also calculated in order to identify those tests that best distinguished between the two groups. Hinting Task appear to be the most sensitive followed by C and I, Strange Stories, Faux Pas, and RMET, in that order. Only RMET did not distinguish between the groups (Table 2) and was not considered for further analysis.

Group Comparisons of Executive Measures
Student t-tests or ANCOVA controlling for IQ and education, were used to compare the performance of the two groups on the 24 measures derived from the eight EF tests. Only the significant results are presented in Table 3.

Performance of the clinical group was impaired compared with the non-clinical group, on 8 out of 24 measures derived from the eight EF tests. The effect size was largest for TOL planning time (tested by \( t \) for unequal variances, effect size using pooled variance \( d = 1.01 \) and using control variance \( d = 1.13 \)) and TOL number of moves \( (\eta_p^2 = 43.4\%) \) followed by Hayling inhibition errors \( (\eta_p^2 = 38.8\%) \) and by Zoo Map execution time \( (\eta_p^2 = 24.3\%) \).

Relation between Performance on EF and ToM Measures
Correlations between ToM variables and EF variables were examined with Education and FSIQ as covariates where appropriate, first in patients and then in controls. Transformed versions were used, when indicated (see Tables 2 and 3). Only EF variables that have at least one significant correlation with a ToM variable are reported in Table 4.

Table 5 presents the percentage of variance \( (r^2) \) of each ToM measure that is explained by the scores on the respective EF tests.

There is no overlap between the best EF and ToM measures within groups. The shared variance differs between tests and groups, ranging from 29.6 to 44.6% in the clinical group and from 15.8 to 24.9% in the non-clinical group, which leaves 75.1 to 84.2% of the variance unaccounted for.

DISCUSSION
We first predicted that some ToM measures would be more sensitive and better discriminate between the pScz and non-clinical group and this hypothesis was confirmed. Hinting Task is the most sensitive \( [\eta_p^2 (\%) = 64.6\%] \) while RMET does not distinguish between the two groups. Hinting Task measures indirect speech acts, requiring distinguishing between explicit, literal, unambiguous content and the intended, implicit, ambiguous content (Lukas, 2011; Bliksted et al., 2014). It requires a sharing of information, some knowledge of the conventions of conversation, the ability to infer the non-literal primary directive component of speech, i.e., the ability to identify and decode the attempt by the speaker to get the listener to do something (Searle, 1975; Hagoort and Indefrey, 2014). The complexity of the task or any of these requirements may explain the sensitivity of this measure. The other tests, C and I, Strange Stories, and Faux Pas are progressively less sensitive in distinguishing between the pScz and non-clinical groups.

In contrast to our prediction but in agreement with Chung et al. (2008), few of the measures used in the present study were sensitive to the pathology even after controlling for IQ. Of the 24 measures derived from the eight EF tests, only eight are significant. The most sensitive of these measures was Tower of London planning time. The clinical group took less time to plan their moves and as a consequence made more moves than the non-clinical group before finding a solution to the problem, although they made a comparable number of errors. However, when they took as much time as the non-clinical group to plan their moves on another task (Zoo Map) \( [\eta_p^2 (\%) = 0.6\%] \), it took them significantly longer to find the solution.
[\eta^2_p (\%) = 24.3] and they made significantly more errors than the non-clinical group [\eta^2_p (\%) = 10.7].

Finally, we predicted that correlations between the different EF measures and ToM tests, would differ sufficiently within and between groups to support the contention that there is more than one component to ToM. As predicted each ToM test had its own best EF correlate. These findings as well as the differences between best EF predictor of ToM in the pScz and non-clinical groups would tend to support the contention that there is more than one ToM component.

A secondary goal of the study of ToM in clinical populations should be to help elucidate the processes involved in the pathology. To this end, and based on the content of the tests, we can derive the following composite image. Patients with paranoid schizophrenia have problems with on-line planning and anticipation of the consequences of their actions (TOL errors) and it is harder for them to switch from one mode of responding to another (Stroop Condition 4 – flexibility time; see also Ibáñez et al., 2014). They especially take more time to plan when confronted with a complex task that requires a lot of thought and planning before initiating any action (Tower of London planning time). They also are more likely to abandon any effort to deduce a rule when it changes without notice (Brixton).

At a social cognitive level, they have difficulties (1) correctly interpreting another's state of mind in order to best be able to explain what one might consider unusual behavior in the context; (2) correctly perceiving and interpreting indirect messages; (3) detecting and understanding an inadvertent, inappropriate comment and the effect that this comment could provoke. These difficulties overlap well with the list of ToM abilities described by Lysaker et al. (2008).

If one is inclined to agree that there is more than one component to 'ToM then what is needed is a model of these components rather than a list. Such a model should be based on the developmental trajectory of the components, the dissociation of neural pathways, and the link between components at different stages of development (Tager-Flusberg and Sullivan, 2000; e.g., are first order beliefs a prerequisite to the development of second order beliefs?). One such model could be as follows: first → second order beliefs ↔ (NB – indicating the possibility of an overlap. See for example Weiner et al., 2012) emotional beliefs, → order beliefs of intention (see for example Baron-Cohen et al., 1999; Brüne et al., 2007; Martino et al., 2007; Zalla et al., 2009).

**REFERENCES**

Abu-Akel, A., and Abusha’leh, K. (2004). ‘Theory of mind’ in violent and nonviolent patients with paranoid schizophrenia. Schizophr. Res. 69, 45–53. doi: 10.1016/S0920-9964(03)00049-5

Abu-Akel, A., and Shamas-Tsoory, S. (2011). Neuroanatomical and neurochemical bases of theory of mind. Neuropsychologia 49, 2971–2984. doi: 10.1016/j.neuropsychologia.2011.07.012

Achim, A. M., Maziade, M., Raymond, É, Olivier, D., Mérette, C., and Roy, M.-A. (2013). Impact of social anxiety on social cognition and function in patients with recent-onset schizophrenia spectrum disorders. Schizophr. Res. 145, 75–81. doi: 10.1016/j.schres.2013.01.012

American Psychiatric Association [APA] (1994). Diagnostic and Statistical Manual of Mental Disorders, Fourth Edn. Washington, DC: American Psychiatric Association.

Baron-Cohen, S., O’Riordan, M., Stone, V., Jones, R., and Plaisted, K. (1999). A new test of social sensitivity: detection of faux pas in normal children and children with Asperger syndrome. J. Autism. Dev. Disord. 29, 407–418. doi: 10.1023/A:1023033012436

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The “reading the mind in the eyes” test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. J. Child Psychol. Psychiatry 42, 241–251. doi: 10.1111/1469-7610.00715

**LIMITATIONS**

The results of this study provide evidence for a multicomponent model of ToM and a message to researchers for the need to identify what these components might be and how they may be affected in various clinical populations. However, there is an important need for replication studies given the relatively small sample and the fact that the results of this study, taken at face value, could be interpreted as being attributable to chance (24 comparisons between two groups, eight significant results – Table 3). However, it should be noted that seven of the eight measures were derived from just three out of the eight tests: Wisconsin Card Sorting Test (3), Zoo Map (2), Tower of London (2). If the distribution was random one would expect the results from more EF tests to be significant. Also, chance remains a viable alternative explanation for most of the correlations in Table 4 (four ToM measures, 24 EF measures and two groups: 11 significant results, 4 of which were at the 0.01 level). These results remain to be confirmed by others.

Finally, the differences between the results on the EF measures and the correlations between the measures and ToM measures may be attributable to the limited psychometric qualities of these tests (see Green et al., 2008) and the fact that we did not control for anxiety (Lysaker et al., 2010; Achim et al., 2011, 2013). There is also a lack of important information concerning the psychometric qualities of each ToM test (Green et al., 2008) although concomitant validity in terms of predicting group membership appears to be acceptable, as does test–retest reliability. Finally, the number of EF and ToM measures used in this study while numerous, do not account for all tests and measures used in this domain. It is quite possible that the inclusion of other tests would modify the findings. More participants from this clinical group need to be tested on this and similar batteries in order to determine the reliability and validity of the results and the applicability of the clinical description to the population.

**ACKNOWLEDGMENT**

The authors would like to thank Dr. Tania Pampoulvat MD, for verifying each PANSS score.
Bora, E., Brüne, M., Heien, T., Veznedaroglu, B. (2006). Theory of mind and executive functions in schizophrenia: mental state decoding may be a better predictor of social functioning than mental state reasoning. *Psychiatry Res.*, 145, 95–103. doi:10.1016/j.psychres.2005.11.003

Bora, E., Sehitoglu, G., Aslier, M., Atabay, I., and Veznedarogly, B. (2006). Deficits of social cognition and neurocognitive deficits in first-episode schizophrenia. *Schizophr. Res.*, 135, 9–17. doi:10.1016/j.schres.2014.01.010

Bora, E., Vucel, M., and Pantelis, C. (2009b). Theory of mind impairment: a meta-analysis. *Schizophr. Res.*, 109, 1–9. doi:10.1016/j.schres.2008.12.020

Bora, E., Yűcel, M., and Pantelis, C. (2009b). Theory of mind impairment: a meta-analysis. *Schizophr. Res.*, 109, 1–9. doi:10.1016/j.schres.2008.12.020

Bora, E., Yücel, M., and Pantelis, C. (2009a). Theory of mind impairment in schizophrenia: the impact of cognitive impairment. *Schizophr. Res.*, 130, 130–136. doi:10.1016/j.schres.2011.04.025

Brüne, M., Abdel-Hamid, M., Lehmkämper, C., and Sonntag, C. (2007). Mental state attribution, mentalistic functioning, and psychopathology: What predicts poor social competence in schizophrenia best? *Schizophr. Res.*, 92, 151–159. doi:10.1016/j.schres.2007.01.006

Burgress, P. W., and Shalllice, T. (1997). *The Haying and Bixton tests*. London: Harcourt Assessment.

Champagne, M., Stip, E., and Joanette, Y. (2005). Cognitive determinants of pragmatic deficits in right brain damaged and schizophrenic individuals: a comparative study. *Brain Cogn.* 5:278.

Champagne-Lavau, M., Stip, E., and Joanette, Y. (2006). Social cognition deficit in schizophrenia: accounting for pragmatic deficits in communication abilities? *Curr. Psychiatry Res.*, 2, 309–315. doi:10.1016/j.cprnec.2010.07.046

Champagne-Lavau, M., Stip, E., and Joanette, Y. (2007). Language functions in right-hemisphere damage and schizophrenia: apparently similar pragmatic deficits may hide profound differences. *Brain* 130:e67. doi:10.1093/brain/aw331

Channon, S., and Crawford, S. (2000). The effects of anterior lesions on performance on a story comprehension test: left anterior impairment on a theory of mind-type task. *Neuropsychologia* 38, 1006–1017. doi:10.1016/S0028-3932(99)00154-2

Chung, Y. S., Barch, D., and Strube, M. (2014). A met-analysis of mentalizing impairments in adults with schizophrenia and autism spectrum disorder. *Schizophr. Bull.*, 40, 602–616. doi:10.1093/schbul/sbt048

Chung, Y. S., Kang, D.-H., Shin, N. Y., Yoo, S. Y., and Kwon, J. S. (2008). Deficit of theory of mind in individuals at ultra-high-risk for schizophrenia. *Schizophr. Res.*, 99, 111–118. doi:10.1016/j.schres.2007.11.012

Corcoran, R., Mercer, G., and Frith, C. D. (1995). Schizophrenia symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. *Schizophr. Res.*, 17, 5–13. doi:10.1016/0920-9964(95)00024-G

Craig, J. S., Hutton, C., Craig, F. B., and Bentall, R. P. (2004). Persecutory beliefs, attributions and theory of mind: comparison of patients with paranoid delusions. Asperger's syndrome and healthy controls. *Schizophr. Res.*, 69, 29–33.
Scherzer, P., Leveillé, E., Achim, A., Boisseau, E., and Stip, E. (2012). A study of SyntaxSemant.

Searle, J. R. (1975). Indirect speech acts. Syntax Semant. 3, 59–82.

Shamay-Tsoory, S. G., Shur, S., Barcai-Goodman, L., Medlovich, S., Harari, H., and Levkovitz, Y. (2007). Dissociation of cognitive from affective components of theory of mind in schizophrenia. Psychiatry Res. 149, 11–23. doi: 10.1016/j.psychres.2005.10.018

Shamay-Tsoory, S. G., Tibi-Elhanany, Y., and Aharon-Peretz, P. (2006). The ventromedial prefrontal cortex is involved in understanding affective but not cognitive theory of mind stories. Soc. Neurosci. 1, 149–166. doi: 10.1080/17470910600985589

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., and Aharon-Peretz, J. (2005). Impaired “affective theory of mind” is associated with right ventromedial prefrontal damage. Cogn. Behav. Neurol. 18, 55–67. doi: 10.1097/01.wbn.0000152228.90129.99

Stone, V., Baron-Cohen, S., and Knight, R. T. (1998). Frontal lobe contributions to theory of mind. J. Cogn. Neurosci. 10, 640–656. doi: 10.1162/08989299898524942

Tager-Flusberg, H., and Sullivan, K. (2000). A componential view of theory of mind: evidence from Williams syndrome. Cognition 76, 59–89. doi: 10.1016/s0010-0277(00)00069-x

Uhlhaas, P. J., Phillips, W. A., Schenkel, L. S., and Silverstein, S. M. (2006). Theory of mind and perceptual context-processing in schizophrenia. Cogn. Neuropsychiatr. 11, 416–436. doi: 10.1080/13546800444000272

Wechsler, D. (1997). The Wechsler Adult Intelligence Scale. Third Edn. San Antonio, TX: The Psychological Corporation.

Weinberger, A. A., Sallquist, J., and Bolnick, R. R. (2012). Young children’s emotion comprehension and theory of mind understanding. Early Educ. Dev. 23, 280–301. doi: 10.1080/10409289.2010.517694

Wilson, B. A., Alderman, N., Burgess, P. W., Emsley, H., and Evans, J. J. (1996). Behavioural Assessment of Dysexecutive Syndrome. London: Harcourt Assessment.

Zalla, T., Sav, A.-M., Stopin, A., Ahade, S., and Leboyer, M. (2009). Faux pas detection and intentional action in Asperger syndrome. A replication on a French sample. J. Autism Dev. Disord. 39, 373–382. doi: 10.1007/s10803-008-0634-y

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Scherzer, Achim, Léveillé, Boisseau and Stip. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.