Clinical Psychology Review 50 (2016) 108–137

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

Deficits in social cognition in first episode psychosis: A review of the literature

Kristin M. Healey [a,*], Cali F. Bartholomeusz [b,c], David L. Penn [a,d]

[a] Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States
[b] Orygen, The National Centre of Excellence in Youth Mental Health, Parkville, Victoria, Australia
[c] The Centre for Youth Mental Health, The University of Melbourne, Parkville, Victoria, Australia
[d] School of Psychology, Australian Catholic University, Melbourne, VIC, Australia

HIGHLIGHTS

- First episode psychosis (FEP) individuals evidence impaired social cognition (SC).
- FEP deficits in emotion perception (EP) and theory of mind (ToM).
- Less research concerning FEP social perception (SP) and attributional style (AS).
- Individuals with FEP and established schizophrenia show comparable SC deficits.
- SC deficits in FEP appear to be stable over time, associated with positive/negative symptoms.

ARTICLE INFO

Article history:
Received 7 July 2016
Received in revised form 23 September 2016
Accepted 8 October 2016
Available online 11 October 2016

Keywords:
First episode psychosis
Early psychosis
Social cognition
Schizophrenia
Theory of mind
Emotion processing

ABSTRACT

Objective: Individuals with chronic schizophrenia (SCZ) consistently show impairments in social cognition (SC) that are associated with functional decline, and work suggests that similar associations exist in first-episode psychosis (FEP). The goal of the current article is to review and synthesize the current body of work examining SC in FEP. Secondary aims are to examine the relationship between SC and symptoms, and change in SC over time in FEP.

Design: Literature is reviewed from four key SC domains: emotion processing (EP), theory of mind (ToM), social perception (SP), and attributional style (AS). Targeted searches of PsycINFO and Google Scholar were conducted to identify relevant manuscripts.

Results: Data from 48 relevant studies (6 longitudinal) were reviewed and integrated.

Conclusions: (1) FEP individuals show consistent deficits in SC compared to healthy controls, most consistently in EP (particularly, fear and sadness recognition) and ToM compared to SP and AS, (2) individuals with FEP and SCZ show comparable SC deficits, (3) some evidence indicates SC deficits in FEP are associated with negative and positive symptoms, and (4) SC appears to be stable over time in FEP.

Contents

* Corresponding author at: UNC Chapel Hill, Dept. of Psychology, Davie Hall, CB #3270, Chapel Hill, NC 27599-3270, United States.
E-mail address: kmhealey@email.unc.edu (K.M. Healey).

http://dx.doi.org/10.1016/j.cpr.2016.10.001
0272-7358/Published by Elsevier Ltd.

1. Introduction

Social cognition (SC) is best defined as a set of neurocognitive processes related to understanding, recognizing, processing, and appropriately using social stimuli in one's environment (Adolphs, 2009; Penn, Corrigan, Bentall, Racenstein, and Newman, 1997). Individuals with schizophrenia consistently show impairments in SC (Savla, Vella, Armstrong, Penn, and Twamley, 2013), and such deficits have been consistently linked to poor functioning, including independent living as well as social and work functioning, across various stages of psychotic illness (e.g., Couture, Penn, and Roberts, 2006; Nuechterlein et al., 2004). The extent of overlap between SC and neurocognition has been an area of debate within the literature. However, a recent meta-analysis of 52 studies indicated that SC accounted for significantly more variance in community functioning than neurocognition (Fett, Viechtbauer, Penn, van Os, and Krabbendam, 2011). Social cognition's real world relevance has been well established in the past decade. This has led to the increased study of SC over the past 10 years, particularly in first episode psychosis (FEP). As a result, SC has been an important target for intervention, with recent efforts aiming to enhance early functional recovery among individuals with FEP.

It is useful to first briefly define chronic schizophrenia (SCZ) relative to FEP. Inclusion criteria for classification as SCZ are inconsistent across studies in the present review. For example, Addington, Saeedi, and Addington (2006a) defined chronic schizophrenia as psychosis lasting greater than three years, whereas Comparelli et al. (2013) defined multi-episode SCZ as more than one acute episode. Others have defined chronic schizophrenia as psychosis lasting for greater than five years (Green et al., 2012; Leung, Lee, and Lee, 2011; Pinkham, Penn, Perkins, Graham, and Siegel, 2007; Vohs et al., 2014). Similarly, there is some inconsistency in criteria that is used to define FEP, however, general criteria define FEP as (1) the first presentation to psychiatric services and (2) duration of psychosis to be within 2 years of illness onset. FEP individuals with diagnoses across the schizophrenia spectrum are included in the present review, which results in heterogeneous patient samples (e.g., ranging from brief psychotic disorder to multi-episode schizophrenia). The benefit of studying this earlier phase of illness is that factors such as aging, illness duration, severity of psychosis, chronic hospitalization, effects of long-term medications, and other variables associated with long-term treatment are not as applicable (Mesholam-Gately, Giuliano, Goff, Faraone, and Seidman, 2009).

Although there has been an increase in the study of SC in FEP, gaps exist in knowledge regarding the course of SC dysfunction at different phases of psychotic illness. Such gaps include a lack of clarity regarding when SC deficits manifest in relation to psychotic illness and how SC impairment progresses in the period of time before and after the first episode (e.g., increases, decreases, or stays constant). Several efforts have been made to review the existing literature on SC at various stages of illness, largely using data from cross-sectional designs. Meta-analytic reviews in FEP found significant impairment in facial emotion recognition (Barkl, Lah, Harris, and Williams, 2014) and theory of mind (ToM) (Bora and Pantelis, 2013). Regarding syntheses of SC research in FEP more broadly, Addington and Piskulic (2013) and McCleery, Horan, and Green (2014) reviewed SC impairment in different phases of illness (e.g., Clinical High Risk (CHR), FEP, and first-degree relatives of individuals with schizophrenia).

However, FEP has not yet undergone an updated, rigorous synthesis of the most current literature, involving all four key social cognitive processes, reviewing comparisons with both healthy individuals and those with more established psychoses, incorporating an evaluation of the relationship SC has with symptomatology, and examining SC changes over time. The present review aims to address such gaps in knowledge concerning SC deficits in FEP and discuss possible clinical implications.

1.1. SC domains and measures

SC is a multi-faceted construct consisting of several sub-domains. Domains of interest in the present review were selected in accordance with key SC domains identified by a panel of experts in the schizophrenia-spectrum research field (Pinkham et al., 2014). They include emotion processing (EP; the ability to perceive and appropriately use emotions), theory of mind (ToM; the ability to infer ones own and others' mental states), social perception (SP; ability to decode and interpret social cues in others), and attributional style (AS; ability to explain the causes or make sense of social interactions and events). There is presently no consensus in the field as to which measures best assess each SC domain. As a result, a heterogeneous group of tasks have been administered with significant conceptual overlap and questionable psychometric properties across studies. For a more detailed discussion of issues related to the measurement and validation of tasks of social cognition, please refer to Pinkham et al. (2014). Measures are referenced in Appendices A–E.

2. Method

2.1. Search strategy

The literature search was performed using the research databases PsycINFO and Google Scholar (January 1990 to May 2016). Search terms for FEP included: "first-episode schizophrenia"; "first-episode psychosis"; and "early psychosis". Search terms for SC included: "social cognition"; "emotion processing"; "emotion perception"; "affect perception"; "affect recognition"; "emotion recognition"; "emotion identification"; "prosody"; "theory of mind"; "mentalizing/mentalising"; "attributions"; "attributional style"; "social perception"; "social knowledge"; and "social cue recognition." The reference lists of published reviews, meta-analyses, and all original studies were also reviewed to ensure all relevant publications were identified for inclusion.

Studies were included if they met the following criteria: (1) included patients with FEP, (2) administered a measure of SC as the primary or secondary outcome, (3) included a comparison group (HC or SCZ), and (4) were written in English. Regarding studies with overlapping samples, the study with the most relevant/usable data was included. However, studies with overlapping samples were included if they provided additional data consistent with the aims of the present study (e.g., additional tasks of social cognition, correlations with symptoms, comparison with SCZ group). An exception to the third criterion was made for longitudinal analyses (i.e., these studies were included even if they did not include comparison groups) due to a relative dearth of longitudinal studies (Hill et al., 2008; Horan et al., 2012; Sullivan et al., 2014). Of note, three of the included studies (two unique samples) conducted targeted recruitment of early-onset FEP individuals (FEP onset prior to 20 years of age) (Amminger, Schäfer, Klier, et al., 2012; Amminger, Schäfer, Papageorgiou, et al., 2012; Seiferth et al., 2009).

2.2. First episode psychosis in comparison to healthy controls

2.2.1. Emotion processing

Within the domain of emotion processing, we will refer to identification and discrimination of facial and vocal emotions as a lower-level process, and the management of emotions as a higher-level process. A total of 25 studies (with 24 non-overlapping samples) have assessed lower and higher-level EP abilities in FEP individuals compared to healthy controls (HCs) (Appendix A). Lower-level EP abilities involve the identification of emotions through facial expressions, voices, or a bimodal combination of the two (e.g., video clips). Higher-level EP abilities involve the understanding and managing of emotional content. Such higher-level EP abilities may be collectively defined as the appropriate regulation and use of emotional information.

In studies comparing global (i.e., overall) scores in lower-level emotion identification tasks, the majority showed significant deficits in FEP individuals (Amminger, Schäfer, Papageorgiou, et al., 2012; Bediou et al., 2007; Comparelli et al., 2013; Edwards, Pattison, Jackson, and Wales, 2001; Herbener, Hill, Marvin, and Sweeney, 2005; Lee et al., 2015; Leung et al., 2011; Mazza et al., 2013; Thompson et al., 2012). In studies that reported effect sizes, values were large and ranged from −0.91 to −1.26 (Comparelli et al., 2013; Thompson et al., 2012). A small number of studies suggested that there are no significant differences between HC and FEP groups in lower-level EP (Achim, Ouellet, Roy, and Jackson, 2012; Dean et al., 2013; Reske et al., 2009). However, each of these studies used EP tasks with unknown psychometric properties. Taken together, FEP individuals show fairly consistent deficits in overall emotion perception when compared to HC individuals.

A subgroup of studies investigated deficits in the recognition of specific emotions (e.g., "sadness") in face and voice tasks (11 studies; Allott et al., 2015; Amminger, Schäfer, Papageorgiou, et al., 2012; Comparelli et al., 2013; Daros, Ruocco, Reilly, Harris, and Sweeney, 2014; Edwards et al., 2001; Kucharska-Pietura, David, Masiak, and Phillips, 2005; Leung et al., 2011; Romero-Ferreiro et al., 2016; Seiferth et al., 2009; Williams et al., 2009; Yang et al., 2015). Fear and sad facial emotion recognition deficits were well replicated (Amminger, Schäfer, Papageorgiou, et al., 2012; Comparelli et al., 2013; Kucharska-Pietura et al., 2005; Seiferth et al., 2009). Evidence in support of specific fear deficits emerged even in small samples (e.g., ≤ 30; Allott et al., 2015; Romero-Ferreiro et al., 2016; Yang et al., 2015). Pronounced fear and sadness deficits were indicated by medium to large effect sizes ranging from −0.71 to −1.04 (Amminger, Schäfer, Papageorgiou, et al., 2012; Comparelli et al., 2013).

With greater power from larger samples ($n = 50$), Comparelli et al. (2013) and Kucharska-Pietura et al. (2005) found that this deficit extended beyond fear, to other negative emotions (e.g., disgust, anger). Consistent with the larger body of literature on SCZ, the majority of the significant emotion specific deficits in FEP involved recognition of negative emotions, as positive emotions have been found to be easier to perceive and often result in ceiling effects for both patients and HCs (e.g., Kohler et al., 2003). However, two studies did not show ceiling effects and found that FEP participants had a deficit in both the identification and labeling of intensity of happy faces (Daros et al., 2014; Williams et al., 2009). More challenging intensity recognition tasks may ameliorate ceiling effects and thus better assess happiness perception. To summarize specific facial emotion recognition in FEP, the majority of the evidence points to deficits in perceiving fear and sadness.

Regarding FEP individuals' ability to detect emotion in spoken voice, 3 of 4 studies suggested a deficit in overall prosodic ability (Amminger, Schäfer, Klier, et al., 2012; Kucharska-Pietura et al., 2005; Thompson et al., 2012). Four studies went on to assess the presence of specific deficits at the emotion level and found fairly consistent evidence for specific deficits in the perception of fear and sadness (Amminger, Schäfer, Papageorgiou, et al., 2012; Edwards et al., 2001; Kucharska-Pietura et al., 2005), with one study showing large effect sizes for anger and surprise (Allott et al., 2015). Given the paucity of evidence, it is difficult to draw clear conclusions about overall and specific voice emotion recognition deficits in FEP, though evidence tentatively points to overall deficits and greater impairment in recognition of negatively valenced voices (e.g., fear and sad).

Only three studies have assessed higher-level EP abilities. Significant deficits were found in all three studies, which involved two different tasks probing typical emotional responses to a series of hypothetical vignettes (Green et al., 2012; Mazza et al., 2013; Thompson et al., 2012). However, it should be noted that Green et al.'s (2012) study collapsed clinical samples of SCZ and FEP when conducting post-hoc group effect analyses, thus this is not a true assessment of deficits specific to FEP.

2.2.2. Theory of mind

A total of 17 studies assessed ToM abilities (16 unique samples; Appendix B), with 25 total task comparisons, the majority of which showed significant deficits in FEP individuals as compared to HCs. We have reviewed the literature according to ToM task type/modality: i) verbal ToM tasks that consist of written stories (11 comparisons), ii) visual cartoon/picture sequencing ToM tasks (5 comparisons), iii) silent animated videos depicting shapes interacting (3 comparisons), and iv) other visual ToM tasks using human stimuli (6 comparisons). ToM was further grouped into first-order and second-order abilities. First-order ToM involved simplistic inferences about a single character's intentions/beliefs. Second-order ToM required more complicated inferences about two characters' interactions and reactions to each other in social situations (Achim et al., 2012).

Results from tasks of verbal ToM indicated a well-replicated deficit in FEP individuals (11 of 11 comparisons) compared to HCs. Four different research groups found evidence of FEP deficits within the Verbal ToM group using the Hinting Task, even after controlling for the effect of IQ (Bertrand, Sutton, Achim, Malla, and Lepage, 2007; Bliksted, Fagerlund, Weed, Frith, and Videbech, 2014; Montreuil et al., 2010; Thompson et al., 2012). Regarding more advanced detection of social faux pas, two studies found significantly impaired ability to both understand and recognize faux pas in FEP individuals (Herold et al., 2009; Ho et al., 2015). There is limited evidence regarding the specificity of deficits in first- vs. second-order verbal ToM, as some work indicates first-order ToM is intact in FEP individuals (Achim et al., 2012) and some indicates non-specific deficits across both levels of ToM (Mazza et al., 2012). Further, on a task of ToM story comprehension comprised of both mentalizing and general comprehension control stories, Langdon, Still, Connors, Ward and Catts (2014) found that FEP individuals performed worse than HCs across both types of stories, indicating a non-specific deficit in ToM. While further work is needed to clarify the specificity of ToM deficits, the overall findings of ToM deficits are robust, as all verbal ToM studies considered neurocognitive indices (e.g., verbal ability) in analyses.

Tasks using visual cartoon stimuli showed somewhat consistent results, with 4 of 5 task comparisons showing significant deficits in ToM (Ho et al., 2015; Inoue et al., 2006; Langdon, Connors, Still, Ward & Catts, 2014; Langdon, Still, et al., 2014). Each study involved a different picture sequencing or cartoon task, and all studies statistically controlled for the possible effect of neurocognitive ability. Two studies found no significant differences in first-order ToM and specific second-order ToM deficits in FEP (Ho et al., 2015; Inoue et al., 2006). Langdon, Still, et al. (2014) reported findings regarding control indices (i.e., inferences about physical factors), and found that FEP patients performed worse on socially-oriented ToM indices. However, one study found the opposite result: FEP individuals showed a specific deficit understanding physical jokes, but not mentalizing (ToM) jokes (Thompson et al., 2012). This is inconsistent with the literature on individuals with established SCZ who show specific mentalizing deficits and intact ability to interpret physical jokes (Corcoran, Cahill, and Frith, 1997). It is possible that this null finding may be related to the poor performance of HC participants, rendering it more difficult to detect a significant group effect ($p = 0.08$) (Thompson et al., 2012). Such poor performance in the HC group may speak to the abstract quality of ToM cartoons, which younger populations may struggle with irrespective of psychotic experiences.

Three studies examined non-verbal ToM abilities among FEP individuals using silent animation tasks (SAT, see http://sites.google.com/site/utafrith for examples). Task administration and scoring were similar across studies (Koelkebeck et al., 2010). Each of these studies accounted for different demographic and neurocognitive covariates, including age, ethnicity, education, parental education, verbal memory, and IQ. While the majority of studies found that FEP individuals showed significant deficits in their ability to interpret videos and use ToM appropriate language when describing ToM shape sequences (3 of 3 studies: Bliksted et al., 2014; Koelkebeck et al., 2010; Ventura et al., 2015), findings were no longer significant in 2 of the 3 studies after controlling for the effects of cognition/IQ (Bliksted et al., 2014; Koelkebeck et al., 2010). Further, Ventura et al. (2015) failed to control for cognitive abilities, thus it is difficult to incorporate the findings. Additionally, Ventura et al. (2015) found that FEP individuals showed a significant deficit in the goal-directed sequences control condition, possibly indicating a more basic cognitive deficit. Further research in this area is required to clarify the role of basic cognition in non-verbal ToM.

The final group of studies investigating ToM utilized cues from human faces, voices, and dynamic interactions. This set of studies used two measures: The Eyes Task and the TASIT. Again, there are discrepancies in the measurement and covariation of neurocognitive factors across these studies, with 3 of the 4 studies omitting this potential confound (Couture, Penn, Addington, Woods, and Perkins, 2008; Green et al., 2012; Kettle, O'Brien-Simpson, and Allen, 2008). All three studies that used the Eyes Task found significant deficits in FEP individuals' ability to identify complex mental states through the eyes only (Couture et al., 2008; Kettle et al., 2008; Mazza et al., 2013). Findings from the 2 studies that used the TASIT suggested that FEP individuals have deficits in accurately detecting lies versus sarcasm (Green et al., 2012), as well as sincerity versus sarcasm (Bliksted et al., 2014). Thus, findings across visual human stimuli tasks reveal consistent deficits in FEP individuals.

2.2.3. Social perception

A total of five studies assessed social perception and social knowledge (sometimes referred to as "social intelligence") in FEP compared to HC participants (Appendix C). Nearly all of these studies (4 of 5) found that FEP individuals had significant SP deficits compared to HCs. Similar to other domains, studies of SP did not consistently control for the effect of IQ. Two studies utilizing the Four Factor Scale of Social Intelligence found significant differences between groups when controlling for the effects of IQ (Bertrand et al., 2007) and verbal/working memory (Montreuil et al., 2010). Of note, the Four Factor Scale of Social Intelligence is considered a measure of SP, but there is significant conceptual and measurement overlap with ToM (e.g., cartoon prediction task, missing cartoons test) and EP (e.g., expression grouping subtest).

Two studies utilized psychometrically validated tasks of SP and found significant deficits in FEP (Addington, Saeedi, and Addington, 2006b; Green et al., 2012). However, these comparisons did not take into account the role of neurocognition despite one study finding SP indices to be significantly correlated with neurocognition (Addington et al., 2006b). Achim et al. (2012) administered a task with unknown psychometric properties (Social Knowledge Task) that appeared to measure ToM (i.e., written vignettes with questions probing how an individual would feel in that situation) and found no significant differences between FEP individuals and HCs ($p = 0.14$). This discrepant finding could possibly be due to the unknown psychometric properties of this new task, or the small sample size ($n = 31$) and lower power.

The limited literature to date suggests SP deficits may be present in FEP individuals in comparison to HCs, though further research is needed in this area.

2.2.4. Attributional style

Four studies examined AS in FEP compared to HCs, and findings were mixed (Appendix D). An et al. (2010) found that FEP individuals had significantly greater perceived hostility bias compared to HCs; when presented with ambiguous hypothetical scenarios, individuals with FEP perceived significantly more hostility in the intentions of others. An et al. (2010) did not find any significant group differences in summed indices of intentionality, blame, and anger (blame index), or self-reported likelihood of responding in an aggressive way (aggression index).

Regarding externalizing bias (i.e., "self-serving bias"), So, Tang, and Leung (2015) found that FEP individuals had a greater tendency to attribute positive (and not negative events) to oneself, but other studies did not (Achim, Sutliff, Samson, Montreuil, and Lecomte, 2016; Fornells-Ambrojo and Garety, 2009). Regarding personalizing bias, or the tendency to attribute negative events to others rather than the environment, there is some evidence that indicates FEP individuals engage more in this bias (Fornells-Ambrojo and Garety, 2009), particularly in individuals reporting greater delusional beliefs (Achim et al., 2016).

2.2.5. Summary

To synthesize findings comparing FEP and HC individuals, work investigating EP suggested moderately consistent deficits in: overall lower-level EP, higher-level EP, and specific recognition of sadness and fear. Some work supports deficits extending to the perception of other negative emotions (e.g., disgust, anger). The following findings are mixed and require further research for clarification: overall vocal emotion recognition score, specific emotion level vocal recognition performance, and the relationship between psychotic symptoms and EP deficits.

Regarding ToM, the strongest support for deficits in FEP compared to HC individuals comes from tasks that assess verbal ToM. These studies are methodologically strong, as they have controlled for the role of neurocognition. It is more difficult to draw clear conclusions regarding non-verbal ToM, especially silent animation tasks. Though the research tends to support the presence of non-verbal deficits in FEP groups, this work has been obfuscated by inconsistencies in covariate measurement and selection.

Extant work investigating SP performance in FEP individuals compared to HCs showed near-consistent deficits in the FEP groups across tasks. However, these tasks are cognitively demanding and a majority of the studies did not account for the potential confound of neurocognition despite evidence of significant associations between SP and neurocognition.

It is difficult to draw conclusions about attributional style in FEP individuals, as studies examined different features of attributional style and within those features, findings were equivocal. Generally speaking, FEP groups show increased likelihood of responding to ambiguous situations in a manner that both blames and perceives hostility in other individuals.

Lastly, to synthesize findings across SC domains, FEP individuals evidenced near-consistent deficits compared to HC groups across domains of EP, ToM, and SP. Further work is needed to determine how attributional style manifests in FEP individuals.

2.3. First episode psychosis in comparison to chronic schizophrenia

2.3.1. Emotion processing

A total of nine comparisons assessed lower and higher-level EP abilities in FEP versus SCZ (8 unique samples; Appendix A). Results were inconsistent regarding between group differences in lower-level emotion perception both in face and voice tasks. A subset of studies found that FEP individuals performed significantly better than SCZ individuals on overall lower-level facial EP tasks (Comparelli et al., 2011; Kucharska-Pietura et al., 2005) and a voice emotion recognition task (Kucharska-Pietura et al., 2005). It should be noted, however, that Comparelli et al.'s (2011) significant Mann-Whitney *U* test finding likely would not have withstood Bonferroni correction for multiple testing ($p = 0.047$), which the authors did not conduct. Four studies, on the other hand, found no between-group differences between FEP and SCZ on lower-level EP (i.e., overall facial emotion recognition) (Addington et al., 2006a; Comparelli et al., 2013; Leung et al., 2011; Pinkham et al., 2007). Further, Vohs et al. (2014) found no significant differences between FEP and SCZ groups in performance on the Bell Lysaker Emotion Recognition Task, a bimodal assessment of EP (Bell, Bryson, and Lysaker, 1997).

At the specific emotion level, two studies found evidence of significant deficits in multi-episode SCZ compared to FEP on identification of faces expressing disgust, fear (Kucharska-Pietura et al., 2005), neutral (Romero-Ferreiro et al., 2016) and anger (Kucharska-Pietura et al., 2005; Romero-Ferreiro et al., 2016). SCZ individuals also evidenced voice emotion recognition deficits (all emotions except for fear) in one study (Kucharska-Pietura et al., 2005). An additional study assessed higher-level EP ability in FEP versus SCZ, and found no significant differences (Green et al., 2012). Of note, studies used varied methods of distinguishing FEP from SCZ participants. Inconsistent definitions of patient groups with respect to time and number of episodes may hinder synthesis of findings and conclusions about what group differences mean in the context of illness stage.

2.3.2. Theory of mind

Three studies assessed ToM in FEP compared to SCZ, and all showed comparable poor performance in both patient groups (Green et al., 2012; Mazza et al., 2012; Vohs et al., 2014). Two of these studies tested mentalizing using the TASIT and Eyes Task (i.e. ability to appropriately interpret cues from human faces/interactions), however IQ was not considered in analyses (Green et al., 2012; Vohs et al., 2014). Mazza et al. (2012) found that both clinical groups performed similarly on measures of verbal ToM, and that ToM performance was not significantly associated with IQ. Regarding potential differences in first- and second-order ToM, FEP and SCZ individuals had comparable deficits in a verbal ToM False Belief Task (Mazza et al., 2012). Thus, findings suggest first- and second-order ToM deficits are present at the FEP illness stage and that these impairments are as severe as those observed in later illness stages, which contradicts three previously discussed studies (Achim et al., 2012; Ho et al., 2015; Inoue et al., 2006).

2.3.3. Social perception

Research investigating differences in SP between FEP and SCZ patients is limited. Two studies that assessed three different tasks of SP generally found that deficits were comparable across groups (Addington et al., 2006b; Green et al., 2012). Of note, Green et al.

(2012) found that FEP individuals performed significantly worse compared to SCZ (and CHR) groups. This cross-sectional finding is counter to what might be expected. However, the authors reported that this finding would not withstand correction for multiple comparisons (eta square = 0.025), though such correction was not conducted. Addington et al. (2006b) found no significant differences between early and chronic groups; however, SCZ was defined as having been ill for greater than three years, which may not have been long enough in duration to detect group differences between FEP and SCZ. As such, further work is required to determine whether SP impairment starts early in the course of illness and remains stable.

2.3.4. Attributional style

There were no studies examining attributional style in FEP compared to SCZ.

2.3.5. Summary

Regarding differences in EP between FEP and SCZ, results of current investigations are mixed. Further research and improved definitions of patient samples are needed to assess whether EP deficits are significantly different during different phases of illness. There was no evidence of significant differences in ToM between FEP and SCZ groups, though there were only three studies examining these potential differences. Additionally, existing work indicates preliminary support that FEP individuals have comparable levels of SP deficits as those with SCZ. No studies investigated potential group differences in attributional style.

Lastly, to synthesize findings across SC domains, FEP individuals evidenced comparable performance to SCZ groups in ToM and SP, whereas differences in EP were inconclusive. Again, further work is needed to determine how attributional style manifests in FEP individuals.

2.4. Relationship of social cognition to psychotic symptoms in FEP

2.4.1. Emotion processing

Current findings in FEP are mixed regarding associations between specific psychotic symptom subscales (e.g., negative, positive symptoms) and performance on tasks of lower-level and higher-level EP. In the present review, only a portion of the included studies (11 of 25 studies) conducted associative analyses between EP and symptoms. Of the studies that reported symptom correlations with EP, 60% found associations between EP and psychotic symptoms. Of note, studies that reported non-significant associations investigated overall EP score, rather than emotion level or subscale scores (Amminger, Schäfer, Papageorgiou, et al., 2012; Green et al., 2012; Kucharska-Pietura et al., 2005; Leung et al., 2011; Romero-Ferreiro et al., 2016).

The most consistent relationship found was between lower-level EP and negative symptoms (Edwards et al., 2001; Lee et al., 2015; Vohs et al., 2014). Herbener et al. (2005) found that decreased EP was correlated with greater levels of negative symptoms, but only after one month of antipsychotic treatment. Further, Piskulic and Addington (2011) found that negative symptoms accounted for 20% of the variance in facial emotion recognition. Regarding positive symptoms, one study found that they were associated with an overall index of EP performance (Lee et al., 2015). Interestingly, two studies found that the positive subscale of the PANSS was associated with the recognition of happy faces, such that more acute positive symptoms were associated with decreased accuracy in recognizing happiness (Daros et al., 2014; Yang et al., 2015). Thus, correlations between symptomatology and EP may reveal more informative patterns of association when examined at the individual emotion level.

Green et al. (2012) found that higher-level EP was not significantly associated with any positive or negative global symptom subscales. Other studies that assessed higher-level EP did not report symptom associations (Mazza et al., 2013; Thompson et al., 2012).

2.4.2. Theory of mind

Correlational analyses between ToM performance and indices of symptomatology have produced inconsistent findings in FEP. The majority of the included studies (13 of 17 studies) reported results from symptom analyses. Eight of these studies found no significant associations between ToM and psychotic symptoms (Bertrand et al., 2007; Bliksted et al., 2014; Couture et al., 2008; Green et al., 2012; Herold et al., 2009; Ho et al., 2015; Inoue et al., 2006; Montreuil et al., 2010). Similarly to EP, of all of the psychotic symptom clusters, the strongest support was for a relationship between ToM and negative symptoms (4 studies; Langdon, Still, et al., 2014; Mazza et al., 2012; Ventura et al., 2015; Vohs et al., 2014). Each of these four studies used a different type of task or a composite index of multiple tasks to assess ToM, thus there was no clear pattern of association across different ToM tasks. Evidence supporting relationships between ToM and other symptom clusters is sparse, though some evidence supports a relationship with positive symptoms (Koelkebeck et al., 2010; Ventura et al., 2015) and disorganized symptoms (Vohs et al., 2014).

2.4.3. Social perception

The current literature shows no consistent pattern of association between psychotic symptoms and SP. Of the four studies that reported such associations, two found that SP was not significantly associated with any symptom subscale (Bertrand et al., 2007; Montreuil et al., 2010), two found that SP was significantly associated with negative symptoms (Addington et al., 2006b; Green et al., 2012), and one found that SP was significantly associated with positive symptoms (Addington et al., 2006b). However, Addington et al.'s (2006b) correlational analyses were derived from a collapsed clinical group of chronic and early psychosis, thus these findings are not specific to individuals with FEP.

2.4.4. Attributional style

Regarding the relationship between symptoms and AS in FEP groups, the Ambiguous Intentions Hostility Questionnaire (AIHQ) hostility index was significantly correlated with the paranoia subscale, but not with other subscales from the Positive and Negative Syndrome Scale (PANSS; Kay, Fiszbein, and Opfer, 1987) (An et al., 2010). Externalizing and personalizing biases were significantly associated with severity of delusional thinking and distress using the Peters Delusion Inventory (So et al., 2015), but was not associated with PANSS scores (Achim et al., 2016).

2.4.5. Summary

Associations between SC deficits and psychotic symptoms in FEP are mixed based on the present review. Of the studies that reported associations, approximately half in each domain (ToM, EP, and SP) found no relationship between SC deficits and psychotic symptoms (i.e., overall scores and subdomains). The remaining half of the studies in each SC domain detected significant associations between negative symptoms and to a lesser degree, positive symptoms. This was consistent with results from a meta-analysis of associations between symptoms and SC in SCZ (Ventura, Wood, and Helleman, 2013).

2.5. Longitudinal studies in FEP

2.5.1. Emotion processing

To date, only two studies have assessed the stability of lower-level EP deficits over time in FEP (Addington et al., 2006a; Hill et al., 2008). Each found that deficits in facial recognition and discrimination tasks were stable, with re-test assessments occurring at three months (Hill et al., 2008) and one year (Addington et al., 2006a). Stability of deficits was irrespective of changes in symptoms, both positive and negative. Horan et al. (2012) investigated the stability of higher-level EP deficits in FEP, finding that MSCEIT scores remained stable over one year. Of note, only one study compared FEP and HC performance longitudinally, and found that the HC group (not the FEP group) evidenced statistically significant improvement on tasks of facial emotion recognition and discrimination (Addington et al., 2006a). Thus, tasks assessing both lower and higher-level EP tend to show stability in FEP individuals.

2.5.2. Theory of mind

Three studies assessed the stability of ToM deficits in FEP individuals over time (Horan et al., 2012; Sullivan et al., 2014; Ventura et al., 2015) (Appendix E). Evidence suggests that indices of ToM tend to be fairly stable (i.e., between 6 months to 1 year follow up across studies). No investigations have assessed the longitudinal stability of ToM performance in FEP compared to HC groups. Thus, the stability of these ToM tasks in typically developing individuals is unknown.

2.5.3. Social perception

Two studies investigated the stability of SP deficits in FEP over a one-year period and both found SP performance was stably impaired (Addington et al., 2006b; Horan et al., 2012). Addington et al. (2006b) noted that this deficit remained stable despite significant improvement in positive symptoms. Similar to Addington et al.'s (2006a) findings with EP, the HC group evidenced significant improvement in SP scores over time while the FEP group did not. Individuals in this sample were in their early-mid 20′s, thus it may be that the FEP group is failing to display normal developmental improvements on this task. Future research in healthy populations should map the trajectories of social cognitive performance during this critical and dynamic neurodevelopmental window.

2.5.4. Attributional style

No studies longitudinally examined attributional style in FEP.

2.5.5. Summary

Though there are relatively few studies investigating longitudinal SC performance following the onset of psychosis, extant work indicated that in those domains that have been examined, SC deficits are stable over time in FEP individuals.

3. Discussion

The current review examined the existing literature on deficits in SC domains (EP, ToM, SP, and AS) in FEP relative to HC and SCZ individuals, examined the stability of SC deficits in FEP over time, and examined relationships between SC domains and psychotic symptoms. There are currently 48 articles included in the review. We will now discuss the findings of this review within the context of a clinical staging model, highlight the implications for treatment, and outline challenges for future research.

3.1. Comparisons among healthy controls and early and late stages of psychotic illness

3.1.1. Emotion processing

The present review suggests EP deficits are present early on in psychotic illness, with specific deficits in interpreting negative emotions (i.e., the recognition of fear and sad facial and vocal expressions). The fact that EP deficits present early, and (although not reviewed herein) are evident prior to the onset of frank psychotic symptoms (e.g., Addington, Penn, Woods, Addington, and Perkins, 2008), supports the notion that EP may serve as a trait-level vulnerability factor for psychosis.

Despite common deficits across early illness stage groups in other SC domains, evidence suggests that there may be differences between SCZ and FEP individuals specifically in lower-level face and voice emotion recognition (Comparelli et al., 2011; Kucharska-Pietura et al., 2005; Romero-Ferreiro et al., 2016). These studies were well-controlled, used well-validated measures and well-defined patient groups, and found a strong group difference in face emotion recognition at the overall EP and specific emotion level between FEP and SCZ, indicative of lower level EP being more impaired in the later phase of illness. This suggests that the FEP stage may be the most optimal time to intervene with targeted SC training to prevent further EP decline. In contrast to these findings, meta-analytic results comparing patient groups to HCs indicated comparable effect sizes between FEP and SCZ individuals on lower-level EP (FEP: $d = -0.88$, Barkl et al., 2014; SCZ: $g = 0.89$, Savla et al., 2013); however, the investigators did not directly compare FEP to the SCZ group. Despite this, Savla et al. (2013) examined moderator variables of EP deficits in SCZ, and found that patients with longer duration of illness had greater deficits (Savla et al., 2013). Similarly, Chan et al.'s (2010) meta-regression analysis found a trend for duration of illness ($p = 0.06$) to be a moderator of facial affect recognition. While duration of illness is conflated with illness chronicity related variables (e.g., prolonged use of anti-psychotic medications), taken together, there is a possibility that EP deficits may worsen over the course of illness.

The present review highlights the need to broaden and improve the measurement of positive EP in particular, and also the perception of non-universal emotions (e.g., alluring). Positive emotions in current tasks of lower-level EP are hampered by ceiling effects, with both patient and control groups performing at near perfect levels. Possibilities for future lines of research could involve the recognition of low arousal positive emotions (e.g., relief, tranquility) in facial and prosodic stimuli. Further, EP tasks may also be improved by developing updated bimodal tasks (e.g., simultaneously presented face/voice) to better emulate real life social interactions.

3.1.2. Theory of mind

Several studies have shown that FEP individuals have deficits in ToM when compared to matched HCs. The strongest and most consistent evidence is from verbal tasks where researchers controlled for the effect of basic cognition. Perhaps this verbal ToM deficit is the first subcategory of ToM to decline (compared to visual/non-verbal), as it is thought to be the more sophisticated higher-order form of ToM (Bora and Murray, 2014). This is consistent with Brüne's (2005) developmental model, which posits that ToM decline occurs in the reverse order of which abilities were acquired (i.e., first second-order ToM deteriorates, then first-order ToM). Similarly, this is in line with the broad neurocognitive evidence showing verbal skills, namely verbal learning and memory, are consistently and significantly impaired early in the course of psychotic illness, more so than other neurocognitive domains (Mesholam-Gately et al., 2009). Though limited, studies included in the present review support Brüne (2005), as first-order but not second-order mentalizing appear to be intact in FEP.

Future work investigating the longitudinal course of first- and second-order ToM abilities would shed light on the pattern of deterioration of such abilities. The underlying neural mechanism that drive ToM deficits should also be explored further in this regard, as one theory posits that in psychosis individuals *hyper*mentalise, or over-attribute, the cause of events to social aspects (Brüne et al., 2011; Walter et al., 2009).

3.1.3. Social perception

Associations between SP deficits and poor functional outcome are well characterized in SCZ (Couture et al., 2006; Fett et al., 2011; Savla et al., 2013; Sergi et al., 2007; Vauth, Rüsch, Wirtz, and Corrigan, 2004). More specifically, SP has shown significant associations with community functioning, social problem solving, and social skills (see review and meta-analysis: Couture et al., 2006; Fett et al., 2011). Thus, it is surprising that SP has received significantly less attention in FEP. While 25 studies investigated EP in FEP compared to HCs, only five studies explored SP. It is likely that SP is less studied because of the limited number of measures (Pinkham et al., 2014). It is possible that few measures exist because of the complexity of the construct and associated difficulty in isolating SP from other social cognitive abilities (e.g., ToM).

Albeit limited, current work suggests nearly consistent significant SP differences in FEP compared to HC groups. Level of SP performance has implications for real world functioning, as social cue perception was found to mediate the relationship between cognition and social functioning in FEP (Addington et al., 2006b). However, conclusions related to SP should be considered with caution, as the literature exploring this domain in FEP is in a nascent stage.

3.1.4. Attributional style

To date, the literature on AS in FEP is very limited. While a meta-analytic study comparing AS in SCZ and HC samples found no significant differences, very few studies were included ($k = 5$, Savla et al., 2013).

Less work has examined AS in FEP. Some research in SCZ, however, indicates that AS is likely a separable construct from SC, as measures of AS often load on a factor that is minimally correlated with other SC factors (e.g., Buck, Healey, Gagen, Roberts, and Penn, 2015; Mancuso, Horan, Kern, and Green, 2011; Van Hooren et al., 2008). Further, evidence shows that AS is associated with paranoid traits across diagnoses (i.e., non-psychotic disorders), rather than discrete psychotic disorders (Savla et al., 2013). This suggests AS may more closely resemble a thinking style associated with personality rather than a performance based SC skill. Subgroups of individuals in early phases of illness may be characterized by a paranoid or hostile cognitive style. This cognitive style may serve as a risk factor for the development of delusional belief systems (Garety and Freeman, 2013). It is possible that particularly maladaptive AS formed at a young age puts individuals at greater risk for later development of psychotic spectrum illness. The question becomes whether such a firmly rooted cognitive style might be responsive to psychosocial training and other treatments. This remains to be answered, as the longitudinal course of AS across early and later stages of psychotic illness is currently unknown.

3.1.5. Relationship of social cognition to psychotic symptoms in FEP

Negative symptoms were related to SC across domains. SC deficits may be associated with negative symptoms because individuals with low SC abilities might be less likely to interact with others effectively, which may in turn increase negative symptoms (e.g., encourage withdrawal and negative affect). Investigating associations between specific negative symptoms and SC indices may help explicate this relationship. Piskulic and Addington (2011) found that while many negative symptoms were associated with SC tasks (i.e., blunted affect, difficulty in abstract thinking, active social avoidance, and lack of spontaneity and flow), stereotyped thinking was the symptom most closely associated with EP/SP (i.e., a form of cognitive inflexibility). Taken together, this suggests that negative symptoms conceptually related to higher-level cognitive processes may be most closely associated with SC ability.

Given that deficits in EP, AS, and ToM have been suggested in prior work to be processes that underlie paranoia in SCZ, significant associations between positive symptoms and such domains would be expected in FEP (Couture et al., 2006). Indeed, we found seven FEP studies that reported significant relationships between positive symptoms and SC performance (2 EP, 2 ToM, 1 SP, 2 AS). However, we did not find support for a relationship between disorganized symptoms and SC in FEP, which is counter to prior work showing moderate associations between SC deficits and disorganized symptoms in SCZ (See meta-analysis, Ventura et al., 2013). It is possible that this lack of a relationship between disorganized symptoms and SC may be a result of FEP samples displaying limited variance in disorganization ratings, on account of being in a remitted state. Several studies recruited FEP individuals approximately 6 months or more following an acute episode. Such individuals may have been in a period of stability associated with remitted symptoms, and for many studies symptom remission was part of the inclusion criteria. The period following FEP is a time when negative symptoms often persist, but positive and disorganized symptoms are ameliorated with antipsychotic treatment (e.g., Melle et al., 2008).

One possible explanation for the mixed findings could be due to variation in underlying moderating and mediating factors. For example, negative symptoms significantly mediated the relationship between ToM intentionality (i.e., from the SAT task) and role functioning in FEP individuals (Ventura et al., 2015). Further, a limited literature has investigated the directionality and structure of relationships in the broader context of cognition, social cognition, symptoms, and social functioning in FEP individuals. Work using structural equation modeling found that the relationship between basic cognition and social functioning was fully mediated by social cognition in a combined sample of first- and multi-episode individuals (Addington, Girard, Christensen, and Addington, 2010), similar to the relationships modeled in established SCZ (Fett et al., 2011). However, these models did not account for the potential role of psychotic symptomatology in early psychosis. Such research in early psychosis is in a nascent stage and warrants further exploration to determine mechanisms that impact the relationship between SC and symptomatology, and to identify those most proximal to functional ability.

3.1.6. Longitudinal course of social cognition in FEP

Broadly, existing longitudinal work supports the stability of SC deficits following the onset of psychosis. However, it is difficult to interpret these results as the majority of studies (4 of 6) are lacking a HC group. As a result, these studies do not consider other possible confounds of time, such as practice effects and typical developmental changes. Two articles using an overlapping participant sample found that FEP individuals performed stably poor across SC indices, while HCs evidenced improvement in EP and SP (Addington et al., 2006a; Addington et al., 2006b). From a neurodevelopmental perspective, this suggests that the developmental trajectory (i.e. maturation/refinement of SC skills) has been interrupted (i.e. arrested or delayed) due to the onset of psychosis (Pantelis et al., 2009). Thus, perhaps instead of asking whether deficits in SC in FEP are stable over time, we should ask whether the rate or 'slope' of improvement in task performance of FEP individuals is similar to that of their HC counterparts (for discussion, see Pantelis, Wannan, Bartholomeusz, Allott, and McGorry, 2015). However, the above interpretation of Addington et al.'s (2006a, 2006b) data should be viewed in light of the mean age of their sample (25.1 years; SD = 8 years), as some participants would have been past the critical developmental window (generally considered to be age 14–25 years, but continuing up to age 30), and there was only one follow-up assessment, at 1-year post-baseline.

3.1.7. Limitations

This review also highlighted several important methodological issues embedded in the current FEP SC literature, such as measurement of social cognition, operational definitions of FEP groups, reporting of sample characteristics, and consideration of neurocognition as a covariate. The SC measures included in the present review are heterogeneous, often have questionable or unknown psychometric properties, and are often scored differently across studies. Thus, it is very difficult to synthesize and compare findings across studies. Further, there is often conceptual and measurement-related overlap in domains/tasks, which makes it difficult to determine deficits in specific domains (Green et al., 2008). To address these issues, efforts are presently underway to psychometrically validate a battery of expert-nominated measures of SC for use in treatment studies (Kern et al., 2013; Olbert et al., 2013; Pinkham et al., 2014). With a psychometrically validated battery of tasks assessing distinct domains of social cognition, there is hope of better capturing true variance in SC deficits instead of potential error variance.

Another issue facing the field of schizophrenia and SC research and thus, understanding SC in early illness, is the lack of measure standardization across different racial and ethnic groups and lack of racial/ethnic heterogeneity of stimuli. This may artificially inflate cross-race and ethnicity differences in SC performance. Such biases likely result in inaccurate detection of deficits in minority participants. African Americans are three times as likely as Caucasians to receive diagnoses of schizophrenia, thus this issue is diagnostically and clinically relevant (Bresnahan et al., 2007).

Current research addressing the question of deficits in FEP compared to chronic SCZ is methodologically limited by the inconsistent definitions of patient groups (i.e., how many years are required for a chronic SCZ categorization). Individuals with diagnoses across the schizophrenia spectrum were included in the present review, which resulted in heterogeneous patient samples (e.g., ranging from brief psychotic disorder to schizophrenia) that possibly had different degrees of SC deficits. Further, several studies identified individuals as FEP who may have fit better in an established SCZ group. These individuals were ill or on antipsychotic medications for up to five years, had a long duration of untreated psychosis (e.g., 7.5 years in Bliksted et al., 2014), or experienced a number of episodes of psychosis that were not appropriately quantified. To more consistently define FEP, a set of criteria should be reliably applied in participant recruitment across research studies. Lastly, sample characteristics and demographic information were inconsistently reported (e.g., years of education), clinical stability (e.g., medication dosages, duration of illness, number of hospitalizations), duration of untreated psychosis, and symptom ratings.

3.1.8. Treatment implications

In order to help individuals with FEP achieve recovery, it is necessary to investigate key determinants of functional decline, and direct treatment and rehabilitation towards intervention targets that will facilitate improvements in functioning. Given SC's impact on functional outcome, a panel of experts at a National Institute of Mental Health (NIMH) Workshop identified SC as a key target for psychosocial interventions in schizophrenia spectrum illnesses (Green et al., 2008).

In established SCZ, SC training interventions are efficacious in treating SC deficits and improving functioning (for a recent review, see Kurtz, Gagen, Rocha, Machado, and Penn, 2016). The next step is to adapt and trial SC interventions that are tailored for young people with early psychosis (Bartholomeusz et al., 2013). For example, FEP impairment in the perception of fear and sadness may be targeted in video game style play incorporating drill-and-practice facial and vocal affect recognition tasks, as well as social role-plays depicting fear/sadness with corrective feedback. Due to greater neural plasticity in adolescence/early adulthood, SC impairment may be more amenable to treatment if delivered during this early illness stage (Fisher, Loewy, Hardy, Schlosser, and Vinogradov, 2013).

4. Conclusions

To our knowledge, this is the first comprehensive, in-depth review of SC deficits in FEP. Studies show that FEP individuals are impaired in both lower-level and higher-level EP. Findings from ToM verbal tasks are strengthened by their consideration of neurocognition, and evidence consistent deficits among FEP individuals. Findings related to SP and AS consistently evidence deficits as well; however, these domains were represented by few studies. Overall, such deficits were found to be stable over time. It should be noted that of the few longitudinal studies conducted, only a sub-set utilized control groups, thus it is unknown whether changes in SC performance over time (or lack of improvement) in FEP are disparate to that of healthy young people with normal developmental trajectories. Healthy progression and development/refinement of SC skills during young adulthood is not well understood and requires further investigation. Future research should also explore the factor structure of social cognition, the necessity of diagnostic clarity, and the relationships between SC and other variables (e.g., neurocognition) to effectively translate this evidence into practice. Tentative conclusions from the present review support that SC deficits are present early in psychotic illness and are comparable to deficits in established SCZ. Such deficits may pre-date illness onset as trait vulnerability factors, or SC may be among the first skills to decline at the onset of illness. As such, targeted SC treatment interventions are likely of maximal benefit when implemented as early as possible in the course of psychotic illness.

Role of funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Contributors

Kristin M. Healey was responsible for the initial conceptualization of the current review (along with D.L.P.), conducted the literature searches, wrote the first drafts of the manuscript and constructed descriptive tables.

Cali F. Bartholomeusz, Ph.D., provided detailed revisions and comments on the manuscript.

David L. Penn helped conceptualize the scope and focus of the review and provided comments on the final draft of the manuscript.

All authors contributed to and approved the final manuscript.

Conflict of interest

All authors declare that they have no conflicts of interest.

Appendix A. Emotion Processing (EP) in First Episode Psychosis (FEP)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
Addington et al. (2006a)	FEP: currently experiencing FEP; <3 months of treatment SCZ: dx >3 years; >2 admissions to the hospital	50 FEP, MA = 25.6 (8.0) Male: 60% Completed HS: 66% Yrs Ill: NR Clor eq: 380 mg/day Cog: NR 53 SCZ MA = 35.5 (7.2) Male: 72% Completed HS: 71.7% Yrs Ill: NR Clor eq: 715 mg/day IQ: NR Avg. hosp: 5	55 HC, MA = 21.2 (6.1) Male: 60% Completed HS: 72.2 IQ: NR	EP (lower-level)	FEIT FEDT (Kerr & Neale, 1993)	NR	None	FEIT • SCZ/FEP vs. HC, face: patient groups collapsed. FEIT HC > SCZ/FEP at time 1 ($F = 7.3$, $p < 0.001$) and time 2 ($F = 15.5$, $p < 0.0001$) • FEP vs. SCZ, face: n.s. differences; effect sizes comparing SCZ to FEP for FEIT = 0.09 • Neurocognition: significant associations FEIT and cognition at all time points FEDT • SCZ/FEP vs. HC, face: patient groups collapsed. FEDT HC > SCZ/FEP at time 1 ($F = 3.1$, $p < 0.001$) and time 2 ($F = 8.6$, $P < 0.0001$) (df not reported) • FEP vs. SCZ, face: n.s. differences; effect sizes comparing SCZ to FEP for FEDT = 0.04 • Neurocognition: significant associations between FEDT and cognition at all time points
Achim et al. (2012)	FEP: age 18–35; in early stages of a psychotic disorder Dx: SCZ (n = 23), SCZA (n = 2), DD (n = 4), PNOS (n = 2)	31 FEP, MA = 24.9 (4.5) Male: 83.87% Edu Category: Hollingshead's categories; 4.0 (1.1) Yrs Ill: NR Duration of tx: 20.9 months Clor eq: NR Estimated IQ: 100.4 (15.1)	31 HC, MA = 25.2 (4.2) Male: 83.87% Yrs Ed: Hollingshead categories; 3.3 (1.2) IQ: 101.8 (10.5)	ToM EP (lower-level) SP	14 photos from Ekman and Friesen (1976) stimuli	NR	Edu	• FEP vs. HC, face: no significant differences between groups ($t(60) = 0.99$, $p = 0.329$
Allott et al. (2015)		30 FEP, MA = 16.8 (1.4) Male: 60% IQ: 90.8 (10)	30 HC, MA = 15.6 (2.0) Male: 50% IQ: 104.4 (11.1)	EP (lower-level)	FACT (Feinberg et al., 1986)	NR	Age, IQ, PANSS total score	• FEP vs. HC, face: HC > FEP on anger (p = 0.017), disgust (p = 0.033), fear (p = 0.040). No significant differences in sadness (p = 0.052), happiness (p = 0.326), and neutral (p = 0.504). • Effect sizes were moderate for anger, disgust, and fear • Adjustment for confounds, only fear remained significant (p = 0.048) • Ceiling effect for happiness (mean % between 98.1 and 100 across both groups) • FEP vs. HC, voice: HC > FEP on anger (p = 0.001) and surprise (p = 0.003). No significant differences in sadness (p = 0.079), fear (p = 0.862), and neutral (p = 0.179) • Effect sizes were large for anger and surprise

(continued on next page)

Appendix A (*continued*)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
								• No longer significant when controlling for confounds
Amminger, Schäfer, Klier, et al. (2012)	FEP: within 6 months following resolution of FE; dx all SCZ	30 FEP, MA = 16.8 (1.4) Male: 60% Yrs Ed: NR Yrs Ill: NR Clor eq: 245.1 (148.1) mg IQ: NR	30 HC, MA = 15.6 (2.0) Male: 50% Yrs Ed: NR IQ: NR	EP (lower-level)	FACT (Feinberg et al., 1986)	NR	Age, sex, edu	• FEP vs. HC, face: HC > FEP in facial emotion recognition, adjusted mean difference of −7.97 (95% CI −13.39 to −2.56; $p =$ 0.004) • FEP vs. HC, voice: HC > FEP in voice emotion recognition, adjusted mean difference of −11.30 (95% CI −11.30 to −1.48; $p =$ 0.011
Amminger, Schäfer, Papageorgiou, et al. (2012)		Same sample as Amminger, Schäfer, Klier, et al. (2012): 30 FEP, MA = 16.8 (1.4) Male: 60% Yrs Ed: NR Yrs Ill: NR Clor eq: 245.1 (148.1) mg IQ: NR	Same sample as Amminger, Schäfer, Klier, et al. (2012): 30 HC, MA = 15.6 (2.0) Male: 50% Yrs Ed: NR IQ: NR	EP (lower-level)	FACT (Feinberg et al., 1986)	No significant associations between PANSS scores and EP	Age, sex, IQ, psychotic symptoms	• FEP vs. HC, face: HC > FEP on fear/sadness aggregate score ($p = 0.004$), fear ($P =$ 0.033), sadness ($p = 0.014$), but not anger ($p = 0.059$). Anger only reached significance when not controlling for confounders. Surprise and neutral, n.s. • Ceiling effect for happiness (mean % between 95.1 and 100 across both groups). • FEP vs. HC, voice: HC > FEP deficits in anger only ($p = 0.024$). surprise was sig until controlling for covariates. ($p = 0.182$)
Bediou et al. (2007)	FEP: drug-naïve patients experiencing first episode of schizophrenia (FES)	40 FEP, MA = 26.9 (5.4) Male: 100% Yrs Ed: 9.6 (2.1) Yrs Ill: NR Clor eq: NR IQ: NR	30 HC, MA = 24.3 (3.3) Male: 100% Yrs Ed: 16.2 (1.1) IQ: NR	EP (lower-level)	Emotion Facial Expression Recognition (Bediou et al., 2005)	NR	Age, edu	FEP vs. HC, face: HC > FEP ($F(1186) = 38.16$, $p < 0.001$)
Comparelli et al. (2011)	FEP: Patients in their first episode of psychosis Dx: 22 SZF 22 SCZ	44 FEP, MA = 25.3 (7.9) Male: 54.54% Yrs Ed: 12.6 (2.3) Yrs Ill: 0.6 (0.7) Clor eq: NR IQ: NR	86 HC, MA = 29.6 (6.4) Male: 58.14% Yrs Ed: 12.5 (2.6) IQ: nr	EP (lower-level)	74 photos from Ekman (1989) stimuli	NR	None	• FEP vs. HC, face: HC > FEP on both subtests ($p < 0.001$) • FEP vs. SCZ, face: FEP > SCZ on subtest A of emotion identification (U = 623.5, Z = −0.19, ($p = 0.047$), but not subtest B (picking a face that expresses an emotion) (n.s.)
	SCZ: Patients with SCZ, >2 acute episodes	38 SCZ MA = 35.1 (8.1) Male: 63.16% Yrs Ed: 12.7 (3.1) Yrs Ill: 10.4 (6.6) Clor eq: mg/day NR IQ: NR Avg. hosp: NR						
Comparelli et al. (2013)	FEP: Recent onset of FE. Dx: SZF, SCZ	50 FEP, MA = 25.7 (7.7) Male: 76% Yrs Ed: 12.7 (2.6)	86 HC, MA = 29.2 (6.7) Male: 46.51% Yrs Ed: 13.2 (2.9) IQ: 112.8 (5.1)	EP (lower-level)	74 photos from Ekman (1989) stimuli	NR	Sex, age, PANSS positive subscale, PANSS total scores	• FEP vs. HC, face, overall score: HC > FEP in both tasks. Effect sizes: Identification (subtest A) -1.27, effect size for Recognition (subtest B) -1.26 • FEP vs. HC, face, specific emotions: HC > FEP in recognition of sadness, disgust, anger, and

		Yrs Ill: 1.1 (0.7) Clor eq: NR IQ: 94.9 (7.7)							fear (p < 0.001 for all). Happiness and surprise, n.s differences Effect sizes: sadness (−1.04), fear (−0.93), disgust (−0.80), and anger (−0.63) • FEP vs. SCZ, face: n.s. differences
Daros et al. (2014)	FEP: <4 months of cumulative lifetime antipsychotic medications allowed. Dx: SCZ (n = 29), BP I with psychosis (n = 28)	44 SCZ MA = 34.3 (7.9) Male: 65.9% Yrs Ed: 12.9 (2.9) Yrs Ill: 10.4 (6.6) Clor eq: mg/day IQ: 92.1 (4.9) Avg. hosp: NR	24 FEP, MA = 22.6 (5.7) Male: 79.2% IQ: 90.7 (17.4)	32 HC, MA = 25.8 (6.8) Male: 34.4% IQ: 102.7	EP (lower-level)	Penn Emotional Acuity Test (PEAT; Erwin et al., 1992)	Recognition of moderately happy faces was significantly correlated with positive subscale of PANSS (r = 0.41, p = 0.03)	IQ	• FEP vs. HC, face: HC > FEP in recognizing mild and moderately happy faces (p < 0.02) • FEP more likely to underestimate intensity/perceive opposite valence of mild happy (F(1,84) = 14.65, p < 0.001) and moderately happy faces (F(1,81) = 17.39, p < 0.001)
Dean et al. (2013)	FEP: patients, age 18–35, FE episode within 1 year of testing; presence of Melbourne criteria (psychosis for >1 week)		30 FEP, MA = 23.4 (4.6) Male: 66.7% Yrs Ed: NR Yrs Ill: NR Clor eq: 248 mg/day NART: 111.9 (6.8)	31 HC, MA = 26.5 (10.3) Male: 61.3% Yrs Ed: NR NART: 108.1 (4.7)	EP (lower-level)	Facial affect intensity rating test (Calder et al., 2000)	NR	Age, pre-morbid intelligence estimate using reading score (NART)	• FEP vs. HC: no significant differences between FEP and HC groups' accuracy ratings
Edwards et al. (2001)	Testing completed 6 weeks–6 months following resolution of FE (for stability); age of onset between 16 and 30, first treated episode Dx: SCZ		29 FEP MA = 22.31 (4.12) Male: 75.9% Yrs Ed: 10.79 (0.94) Yrs Ill: NR Clor eq: 248.79	24 HC, MA = 21.50 (4.09) Male: 62.5% Yrs Ed: 10.54 WAIS-R: 97.63 (13.26)	EP (lower-level)	FACT (Feinberg et al., 1986)	Negative symptoms significantly correlated with FACT subtests; subtests and statistics NR	None	• FEP vs. HC, face, specific emotions: HC > FEP/OPD on fear (t = 4.15, df = 300, p < 0.001) and sad (t = −2.72, df = 300, p < 0.01) faces • FEP vs. HC, voice, specific emotions: no significant differences between groups on any emotions • IQ made an independent contribution to the

(continued on next page)

Appendix A (*continued*)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
		WAIS-R: 86.83						prediction of emotion recognition ($t = 4.10$, $df = 72$, $p < 0.001$); was not controlled for
Green et al. (2012)	FEP: age 18–45; AAO within 2 years of study participation; most recovering from FE (e.g. approx. 2–3 months after hosp, after med level was stabilized) Dx: SCZ ($n = 46$), SCZA ($n = 10$), SZF ($n = 25$)	81 FEP, MA = 22.02 (4.18) Male: 75% Yrs Ed: 12.50 (1.96) Yrs Ill: NR Clor eq: NR IQ: NR	46 HC MA = 22.20 (3.51) Male: 63% Yrs Ed: 13.86 (1.97) IQ: NR	EP (higher-level) ToM SP	MSCEIT (Mayer et al., 2003)	No significant associations between SANS and SAPS global scores and MSCEIT.	Parental education	• FEP vs. HC, higher-level: HC > FEP, Cohen's $d = 0.76$ • FEP vs. SCZ, higher-level: No significant differences between groups
	SCZ: FE >5 years before testing Dx: SCZ (n = 48), SCZA (n = 5)	53 SCZ MA = 34.77 (7.89) Male: 66% Yrs Ed: 13.96 (1.64) Yrs Ill: NR Clor eq: NR mg/day IQ: NR Avg. hosp: NR	47 SCZ HC MA = 33.02 (5.32) Male: 72% Yrs Ed: 14.45 (1.69) IQ: NR					
Herbener et al. (2005)	FEP: Shortly after experiencing their FEP (acutely ill); <4 weeks treatment; unmedicated for 5 days before testing	13 FEP (No demographic information reported)	13 HC (No demographic information reported)	EP (lower-level)	Penn CNP (Gur et al., 2001)	Following tx, negative sx significantly associated with EP ($r = -0.70$, $p < 0.01$); held controlling for IQ	IQ	• FEP vs. HC, lower-level: HC > FEP, on the emotional acuity task ($F = 24.63(1, 24)$; $p < 0.001$) and emotion differentiation task ($p < 0.002$)
Kucharska-Pietura et al. (2005)	FEP: age 18–60; < or = to 2 episodes of psychosis	50 early stage SCZ, inpatients MA = 23.1 (4.2) Male: 50% Yrs Ed: 12.1 (1.8) Yrs Ill: 2.0 (1.3) Clor eq: 365 (165) IQ: NR	HC MA = 36.8 (13.4) Male: 48% Yrs Ed: 13.5 (3) IQ: NR	EP (lower-level)	FERT VERT (Kucharska-Pietura et al., 2003)	No significant association between PANSS total score and FERT or VERT.	Age, edu, current mood ratings	FERT (Faces) • FEP vs. HC, lower-level, face, overall: HC > FEP, overall score mean difference = 13.41; $p < 0.001$. • FEP vs. HC, lower-level, face, specific emotions: HC > FEP across all emotions ($p < 0.05$) • FEP vs. SCZ, lower-level, face: FEP > SCZ, mean difference = 8.33; $p < 0.001$ • FEP vs. SCZ, lower-level, face, specific emotions: FEP > SCZ, in positive valence/-negative valence groups ($p < 0.05$), but not specific emotions
	SCZ: age 18–60, >2 episodes of psychosis	50 SCZ inpatients MA: 41.6 (10.3) Male: 52% Edu: 12.3 (6.6) Yrs Ill: 13.2 (6.6) Clor eq: 383 (131) IQ: NR Avg. hosp: NR						VERT (Voices) • FEP vs. HC, lower-level, voice, specific: HC > FEP, mean difference = 27.94; $p < 0.001$ • FEP vs. HC, lower-level, voice, specific: HC > FEP, across all emotions, $p < 0.01$ • FEP vs. SCZ, lower-level, voice, overall: FEP > SCZ, mean difference = 12.62; $p < 0.001$ • FEP vs. SCZ, lower-level, voice, specific emotions: SCZ more impaired than FEP across emotions ($p < 0.05$), except fear (n.s.)
Lee et al. (2015)	FEP participants were recruited <1.5 months from hospitalization with stabilized medications. 21/24 FEP mild positive symptoms, 3/24 moderate.	24 FEP, MA = 20.5 (3.3) Male: 33.3% Yrs Ed: 13.3	46 HC, MA = 20.8 (3.5) Male: 54.3% Yrs Ed: 13.2 (1.9) IQ: 105.8 (12.5)	EP (lower-level)	Japanese and Caucasian Facial Expressions of Emotion (JACFEE) (Matsumoto & Ekman, 1988)	Significant relationship between perceptual aberration scale (Chapman et al., 1978) ($r = -0.55$, $p < 0.001$)	IQ, Gender	FEP vs. HC, lower level, overall: HC > FEP, $p = 0.023$, Cohen's $d = 1.23$

		(2.3) IQ: 96.0 (15.7)						
Leung et al. (2011)	FEP: Duration of untreated psychosis <2 years; clinically stable; no hospitalizations <6 weeks Dx: SCZ	50 FEP, MA = 20.7 (2.72) Male: 50% Yrs Ed: 14.6 (2.42) Months Ill: 16.86 Age of onset: 19.51 (2.89) Number of admissions: 0.64 (0.67) Clor eq: 258.58 (216.28)	26 HC MA = 21.7 (2.72) Male: 46.2% Yrs Ed: 16.9 (2.71)	EP (lower-level, emotion specific)	Japanese and Caucasian Facial Expressions of Emotion (JACFEE) (Matsumoto & Ekman, 1988)	No significant association between overall score and SANS and SAPS subscores.	Raven's standardized score	• FEP vs. HC, lower-level, overall: HC > FEP, $p > 0.001$ • FEP vs. HC, lower-level, specific emotions: HC > FEP, significant deficit in recognition of surprise, fear and disgust ($p < 0.05$); but not for sadness and anger ($p = 0.72$–0.95). However, only 2 stimuli per emotion category • FEP vs. SCZ, lower-level, overall: No significant differences between groups p$P =$ 0.365) • FEP vs. SCZ, lower-level, specific emotions: no significant differences
	SCZ: diagnosed with SCZ for <5 years; clinically stable; no hospitalizations within 6 weeks Dx: SCZ	51 SCZ MA = 43.5 (8.61) Male: 60.8% Yrs Ed: 12.5 (3.7) Months Ill: 229.41 (116.68) Clor eq: 319.38 (250.74) mg/day Avg. hosp: 3.02 (3.04)	28 HC MA = 44.8 (8.21) Male: 60.7% Yrs Ed: 13.32 (4.87)					
Mazza et al. (2013)	FEP: within 3 months of initial dx (e.g. first time of hospitalization or presentation for services) Dx: SCZ (n = 4), SCZA ($n = 4$), PNOS (n = 3), delirium ($n = 1$)	12 FEP, MA = 27.92 Male: 66.67% Yrs Ed: 13.5 (2.3) Yrs Ill: NR Clor eq: NR IQ: NR	12 HC MA = 27.25 (9.3) Male: 41.67% Yrs Ed: 15 (2.3) IQ: NR	EP (higher-level) ToM	Emotion Attribution Task (Blair & Cipolotti, 2000; Mazza et al., 2007)	NR	None	FEP vs. HC, higher-level: HC > FEP, Emotion Attribution Task ($t(1,22) = -3.26$, $p < 0.01$)
Pinkham et al. (2007)	FEP: individuals early in the course of illness, psychosis <5 years via review of clinical records. Dx: SCZ ($n = 17$), SCZ ($n = 2$), PNOS ($n = 2$)	21 SSI MA = 24.62 (4.92) Male: 66.7% Yrs Ed: 13.71 (2.33) Yrs Ill: 2.0 (1.38) Clor eq: 250.85 (280.17) IQ: NR	21 HC MA = 27.62 (4.31) Male: 47.6% Yrs Ed: 16.67 (2.85) IQ: NR	EP (lower-level)	FEIT FEDT (Kerr & Neale, 1993)	NR	None	FEIT • FEP vs. HC, lower-level: HC > FEP, FEIT ($p = 0.014$) • FEP vs. SCZ, lower-level: No significant differences between groups FEDT • FEP vs. HC, lower-level: HC > FEP, FEDT approached statistical significance ($p = 0.051$) • FEP vs. SCZ, lower-level: No significant differences between groups
	SCZ: psychosis >5 years. Dx: SCZ ($n = 19$), SCZA (n = 9)	28 SCZ MA = 39.57 (8.38) Male: 50% Edu: 14.68 (2.14) Yrs Ill: 17.0 (7.72)						

(continued on next page)

Appendix A (*continued*)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
		Clor eq: 421.73 (271.07) IQ: NR Avg. hosp: NR						
Reske et al. (2009)	FEP: individuals in their first episode of psychosis; no lifetime comorbidities	12 FEP, MA = 31.94 Male: 50% Yrs Ed: 13.78 (3.3) Yrs Ill: NR Clor eq: NR IQ: NR	15 HC MA = 31.94 (6.03) Male: 66% Yrs Ed: 13.11 (3.74) IQ: NR	EP (lower-level)	115 item Subset of stimuli from Erwin (1992)	NR	None	FEP vs. HC, lower-level: trend level difference HC > FEP, ($F(1,23) = 4.06, p = 0.056$)
Romero-Ferreiro et al. (2016)	FEP: individuals 12–18 years old	20 FEP, MA = 15.6 (1.6) Male: 65% Clor eq: 206 (129) 19 multiepisode schizophrenia (MES), MA = 43.9 (9.5) Male: 68%	20 FEP HC MA = 15.2 (1.6) Male: 65% 19 MES HC MA = 42.3 (12.9) Male: 55%	EP (lower-level)	NimStim set (Tottenham et al., 2009)	PANSS depressed factor associated with fear accuracy ($r = 0.473, p = 0.035$)	None	• FEP vs. FEP HC, lower level: HC > FEP in fear ($p = 0.007$). No other emotions were significant (angry, sadness, happiness, neutral) • FEP vs. MES, lower level: FEP > MES for angry and neutral faces only ($p < 0.05$). No other emotions were significantly different (sadness, happiness, fear)
Seiferth et al. (2009)	FEP: individuals <19 years	12 FEP, MA = 17.8 (1.4) Male: 100% Yrs Ed: NR Yrs Ill: NR Clor eq: 231 (111) IQ: NR	12 HC MA = 17.9 (1.5) Male: 100% Yrs Ed: NR IQ: NR	EP (lower-level)	Facial Emotions for Brain Activation (Gur et al., 2002a; Schneider et al. 2006)	NR	None	• FEP vs. HC, lower-level, overall: HC > FEP, ($F(1,21) = 5.95; p = 0.02$) • FEP vs. HC, lower-level, specific: impaired recognition sad, angry, fear (p values NR)
Thompson et al. (2012)	FEP: 1 week of daily psychotic symptoms (Comprehensive Assessment of At Risk Mental States; Yung 2005); <6 months tx for a psychotic disorder	40 FEP, outpatients MA = 20.5 (2.5) Male: 62.5% Yrs Ed: 12.5 (1.6) Yrs Ill: NR Clor eq: NR Meds: 30/40 on Antipsychotics Estimated IQ: 106.4 (10.6) (based on NART)	30 HC MA = 19.3 (2.9) Male: 41.9% Yrs Ed: 13.2 IQ: 104.9 (12.5) WASI	EP (higher-level, lower-level) ToM SP	DANVA-2 (Nowicki & Duke, 1994) MSCEIT (branch 4) (Mayer et al., 203)	NR	Age, sex, IQ estimated based on Reading Test (NART)	DANVA-2 (lower-level) • FEP vs. HC, lower-level, face: HC > FEP, post-hoc Tukey's tests adjusted for multiple testing; $p < 0.05$ • FEP vs. HC, lower-level, voice: HC > FEP, post-hoc Tukey's tests adjusted for multiple testing; $p < 0.01$ • FEP vs. HC: Effect size for combined face/voice, 0.91 MSCEIT (higher-level) • FEP vs. HC, higher-level: HC > FEP, ($p < 0.01$). Effect size, 0.64 • Controls performed below average (100), HC mean = 92.0, SD = 10.38
Vohs et al. (2014)	FEP: age 18–65; <5 years of a documented initial dx of a SZ spectrum disorder; no medication changes or hospitalizations within 30 days Dx: SCZ, SCZA, PNOS	26 FEP, MA = 23.81 (3.63) Male: 81% Yrs Ed: 13 (1.87) Yrs Ill:	Used non-psychiatric control group from substance use disorders program	EP (lower-level) ToM	BLERT (Bell et al., 1997)	BLERT not significantly associated with postive sx ($r = -0.07$); associated with negative sx ($r = -0.49, p < 0.01$) and disorganized sx ($r = -0.66, p < 0.001$)	Age	• FEP vs. SCZ, lower-level: post-hoc tests however revealed no significant differences between FEP and SCZ

	SCZ: age 18–65; >5 years of a documented initial dx of a SZ spectrum disorder; no medication changes or hospitalizations within 30 days Dx: SCZ, SCZA	Clor eq: 307.2 (60.2) IQ: NR 72 SCZ MA = 51.01 (6.94) Male: 88% Yrs Ed: 12.72 (1.81) Yrs Ill: Clor eq: 420.5 (93.4) IQ: NR Avg. hosp: NR						
Williams et al. (2009)		FEP: Within 8 months of first presentation with psychotic symptoms Dx: SCZ	28 FEP, MA = 19.9 (3) Male: 71.4% Yrs Ed: 11.45 (1.7) Yrs Ill: NR Clor eq: 258 (216) Estimated IQ: 95.3 (8.1)	72 HC MA = 20 (2.8) Male: 75% Yrs Ed: 12.1 (2.1) Estimated IQ: 103.5 (6.0)	EP (lower-level)	Emotion Identification Task (Williams, Palmer, & Liddell et al., 2006; Williams et al., 2007)	NR	None
• FEP vs. HC, lower-level, face: HC > FEP on all reported comparisons: fear: $t = -3.60$, $p = 0.001$; happy: $t = -4.21$, $p < 0.001$); neutral $t = -3.23$, $p = 0.002$)								
Yang et al. (2015)	FEP: Psychosis for <3 years, age 16–28, antipsychotic medication for <2 weeks	30 FEP, MA = 22.3 (3.2) Male: 50% Yrs ed: 11.9 (3.2)	30 HC, MA = 24.6 (2.2) Male: 50% Yrs ed: 13.2 (3.5)	EP (lower-level)	176 Ekman and Friesen (1976) stimuli (Ekman & Friesen, 1976) morphed to express emotional degree (10–90%)	PANSS positive subscale predicted accuracy in identifying moderate happy faces in regression analyses ($p < 0.05$)	None	• FEP vs. HC, lower-level, face: HC > FEP on moderate and high degree faces across emotions (happy, disgust, and fear; $p < 0.05$). No significant differences between groups in recognition of low degree happy and fear faces

Confounds include variables that are controlled for in subsequent analyses. ANCOVA = Analysis of Covariance; AS = Attributional Style; CI: Confidence Interval; Clor eq = Chlorpromazine equivalent (mg/d); DANVA-2 = Diagnostic Analysis of Nonverbal Behavior – 2; DD = Delusional Disorder; DUP = Duration of Untreated Psychosis; Dx = Diagnosis; EP = Emotion Processing; FACT = Facial Affect Computer Tasks; FE: First Episode; FEP = First Episode Psychosis; FEIT = Face Emotion Identification Task; FEDT = Face Emotion Discrimination Task; FEP = First Episode Psychosis FERT = Face Emotion Recognition Task; HC = Healthy Control; HS = High school; MA = Mean Age; MSCEIT = Mayer-Salovey-Caruso Emotional Intelligence Test; NART = National Adult Reading Test; NR = Not Reported; n.s. = non-significant; OPD = Other Psychotic Disorder; PANSS = Positive and Negative Syndrome Scale; PNOS = Psychosis NOS; SCZ = Schizophrenia; SCZA = Schizoaffective; SP = Social Perception; SZF = Schizophreniform; Sx = symptoms; ToM = Theory of Mind; Tx = Treatment; VERT = Voice Emotion Recognition Task; VT = Verbal Test; WAIS-R = Wechsler Adult Intelligence Scales-Revised; Yrs Ed = Years of Education; Yrs Ill = Number of Years Ill; WAIS = Wechsler Adult Intelligence Scale.

Appendix B. Theory of Mind (ToM) in First Episode Psychosis (FEP)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
Achim et al. (2012)	FEP: age 18–35; in early stages of a psychotic disorder Dx: SCZ ($n = 23$), SCZA ($n = 2$), DD ($n = 4$), PNOS ($n = 2$)	31 FEP, MA = 24.9 (4.5) Male: 83.87% Edu Category: Hollingshead's categories; 4.0 (1.1) Yrs Ill: NR Duration of tx: 20.9 months Clor eq: NR Estimated IQ: 100.4 (15.1)	31 HC, MA = 25.2 (4.2) Male: 83.87% Yrs Ed: Hollingshead categories; 3.3 (1.2) IQ: 101.8 (10.5)	ToM EP SP	Combined Stories Test (Verbal ToM) (Achim et al., 2012)	NR	Edu	• FEP vs. HC, mentalizing index: HC > FEP ($t(60) = 3.28$, $p = 0.002$) • FEP vs. HC, non-social reasoning index: HC > FEP, $U = 348.0$, $Z = 1.96$, $p = 0.050$) • FEP vs. HC, first-order ToM: N.s.; ceiling effect. FEP ($m = 2.9$, $sd = 0.3$), HC ($m = 2.8$, $sd = 0.4$) (out of possible 3)
Bertrand et al. (2007)	FEP: patients with a FE. Dx: SCZ ($n = 21$), SCZA ($n = 8$), PNOS ($n = 6$), and SZF ($n = 1$)	36 FEP, MA = 22.78 (3.37) Male: 64% Yrs Ed: NR Yrs Ill: NR Days since beginning tx: mean = 143; range 0–529 days DUP (weeks): 62.24 (86.95) range 1–310 Clor eq: NR IQ: 92.31 (15.94)	25 HC MA = 24.19 (3.55) Male: 52% Yrs Ed: NR IQ: 106.31 (10.90)	ToM SP	Hinting Task (Verbal ToM) (Corcoran et al., 1995)	No significant associations	IQ	• FEP vs. HC: HC > FEP, ANCOVA controlling for IQ, ($F(1,57) = 5.221$, $p = 0.03$)
Bliksted et al. (2014)	FEP: age 18–30; met criteria for SCZ; < 3 months of antipsychotic medication prior to dx interview. Patients tested both at home and in clinic	36 FEP, MA = 22.7 (95% CI: 21.6; 23.7) Male: 52.8% Yrs Ed: 12.1 (95% CI: 11.3; 12.9) DUP (yrs): 7.5 (6.0; 9.8) Days of FES dx: 189 (95% CI: 256; 222) Clor eq: NR DART: 27.2 (95% CI: 24.4; 30.0) IQ: 86.4 (95% CI: 81.4; 91.5)	36 HC MA = 22.7 (95% CI: 21.6; 23.7) Male: 52.8% Yrs Ed: 13.4 (95% CI: 12.6; 14.1) DART: 31.8 (95% CI: 30.0; 33.6) IQ: 97.4 (95% CI: 93.0; 101.8)	ToM SP	Hinting Task (Verbal ToM, 2 versions) (Corcoran et al., 1995) TASIT (Visual ToM, human stimuli) (McDonald et al., 2006) Animated Triangles (Abell et al., 2000; Castelli et al., 2000)	No significant associations	Age, Edu, Functional IQ	TASIT • FEP vs. HC, visual human stimuli: HC > FEP, generalized linear modeling used; HC performed significantly better than HC on ToM indices after controlling for covariates: total score ($p < 0.01$), "simple sarcasm" ($p = 0.002$), "paradoxical sarcasm" ($p = 0.04$), "say" ($p < 0.05$), "think" ($p < 0.05$). N.s. differences on subscales "Sincere", "Feel", and "Do" • HC > FEP, sincerity v. sarcasm ($t(70) = -2.17$, $p = 0.03$), patients misidentify sarcasm as sincere Hinting Task • FEP vs. HC, verbal ToM: HC > FEP, ANCOVA controlling for covariates ($p < 0.001$), Corcoran et al. (1995) version • FEP vs. HC, verbal ToM: N.S. ANCOVA controlling for covariates ($p = 0.07$), Marjoram et al. (2005) version Animated Triangles Task

								• FEP vs. HC, visual ToM: HC > FEP on only one index of ToM after controlling for covariates, appropriateness total ($p = 0.01$) All other indices n.s. (Intentionality ToM, Appropriateness ToM, Target item ToM)
Couture et al. (2008)	FEP: < 5 years of a documented initial dx of a SZ spectrum disorder	26 FEP, MA = 24.9 (5.1) Male: 88% Completed HS: 85% Yrs Ill: NR Clor eq: NR IQ: NR	41 HC MA = 23.0 (5.9) Male: 93% Completed HS: 73% IQ: NR	ToM	The Eyes Task (Visual ToM, human stimuli) (Baron-Cohen et al., 2001)	No significant associations	Age, edu, ethnicity	FEP vs. HC: HC > FEP in planned simple contrasts controlling for covariates ($p > 0.05$)
Green et al. (2012)	FEP: age 18–45; AAO within 2 years of study participation; most recovering from FE (e.g. approx 2–3 months after hosp, after med level was stabilized) Dx: SCZ ($n = 46$), SCZA ($n = 10$), SZF ($n = 25$) SCZ: FE >5 years before testing Dx: SCZ ($n = 48$), SCZA ($n = 5$)	81 FEP, MA = 22.02 (4.18) Male: 75% Yrs Ed: 12.50 (1.96) Yrs Ill: NR Clor eq: NR IQ: NR 53 SCZ MA = 34.77 (7.89) Male: 66% Yrs Ed: 13.96 (1.64) Yrs Ill: NR Clor eq: NR mg/day IQ: NR Avg. hosp: NR	46 HC MA = 22.20 (3.51) Male: 63% Yrs Ed: 13.86 (1.97) IQ: NR 47 SCZ HC MA = 33.02 (5.32) Male: 72% Yrs Ed: 14.45 (1.69) IQ: NR	EP (higher-level) ToM SP	TASIT (Visual ToM, human stimuli) (McDonald et al., 2006)	No significant associations	Parental edu	• FEP vs. HC, higher-level: HC > FEP, TASIT total score, Cohen's $d = 1.06$ • FEP vs. SCZ, higher-level: No significant differences between groups
Herold et al. (2009)	FEP: <5 years of psychosis. Dx: SCZ	18 FEP MA = 28.7 (10.3) Male: 61.1% Yrs Ed: NR Yrs Ill: 3.4 (4.1) Clor eq: NR IQ: 99.5 (13)	21 HC MA = 27.4 (6.5) Male: 52.4% Yrs Ed: NR IQ: 115.2 (9.6)	ToM	Faux Pas Task (Verbal ToM) (Abu-Akel & Abushua'leh, 2004) Index: combined understanding/recognizing faux pas score	No significant associations	None	• FEP vs. HC, verbal ToM: HC > FEP ($t = 2.52$, $p = 0.016$), not controlling for significant differences in IQ • Authors created two subgroups matching in IQ ($n = 11$ in each). Removed individuals with IQ scores >120 or <97. N.s. difference in IQ in subgroups HC > FEP subgroups were significantly different with a Kruskal-Wallis non-parametric test for unequal variances ($\chi^2(1,22) = 4.416$, $p = 0.036$)
Ho et al. (2015)	FEP: clinically stable outpatients	41 FEP MA = 27.7 (6.5) Male: 56.1% Yrs Ed: 12.93 (3.3) IQ: 110.3 (11.8) DUP: 10.1 (10.8) months DOI: 20.0 (13.0) months Clor eq: 418.8 (210.8)	42 HC MA = 25.4 (6.9) Male: 52.4% Yrs ed: 13.8 (2.8) IQ: 116.1 (9.9)	ToM	Yoni task (Visual ToM) (Shamay-Tsoory et al., 2007) Faux Pas Task (Verbal ToM) (Stone et al., 1998)	Significant associations between Yoni Task and SANS/SAPS did not survive Bonferroni correction for multiple comparisons. Associations with faux pas task were not reported.	IQ, Letter-Number Span Test (Gold et al., 1997), delayed logical memory (Gong et al., 1989)	• FEP vs. HC, visual ToM, first-order: no significant differences in Bonferroni adjusted pairwise comparisons cognitive ToM ($p = 0.141$), affective ToM ($p = 0.298$), or physica inference ($p = 0.263$) • FEP vs. HC, visual ToM, second-order: significant differences in Bonferroni adjusted pairwise comparisons controlling for IQ/-LNST in affective ToM ($p = 0.0094$) • N.s. differences in cognitive ToM and physical ToM after controlling for IQ • FEP vs. HC, verbal ToM, second order: significant differences in Bonferroni

(continued on next page)

Appendix B (*continued*)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
								adjusted pairwise comparisons controlling for delayed logical memory in faux pas understanding ($p < 0.001$) and recognition ($p = 0.002$) • N.s. control questions
Inoue et al. (2006)	FEP: in remission after FEP; excluded for acute symptoms Dx: SCZ (all paranoid type)	30 FEP, MA = 27.03 (6.07) Male: 16.67% Yrs Ed: 13.77 (1.94) Yrs Ill: 2.78 (1.57) Clor eq: NR IQ: 103.07 (6.34)	30 Controls MA = 27.93 (6.65) Male: 16.67% Yrs Ed: 13.07 (2.08) IQ: 102.03 (15.65)	ToM	Picture arrangement task (Visual ToM) (Inoue, Yamada, Tonooka, & Konba, 2004)	No significant associations	None	• FEP vs. HC, visual ToM, first-order: • First-order false belief, N.s. differences (Fisher's exact test, $p = 0.678$) • FEP vs. HC, visual ToM, second-order: Second-order false belief question, (Fisher's exact test, $p < 0.0001$)
Kettle et al. (2008)	FEP: presently experiencing FE Dx: SCZ ($n = 11$), SZF ($n = 1$), SCZA ($n = 1$)	13 FEP MA = 20.76 (0.89) Male: 84.62% Yrs Ed: 12.38 (0.67) Yrs Ill: NR Clor eq: NR IQ: NR	16 Community HC MA = 19.34 (0.73) Male: 50% Yrs Ed: 12.5 (0.50) IQ: NR 27 University HC MA = 19.17 (0.23) Male: 22.22% Yrs Ed: 12.33 (0.19) IQ: NR	ToM	The Eyes Task (Visual ToM, human stimuli) (Baron-Cohen et al., 2001)	NR	Gender	• FEP vs. HC (community HC): HC > FEP, ($p < 0.033$) • FEP vs. HC (university HC): HC > FEP ($p < 0.001$) Gender: unequal gender distributions, ANOVAS re-run with gender as a factor, no interactions or main effects ($ps > 0.217$)
Koelkebeck et al. (2010)	FEP: Medicated outpatients with FE Dx: SCZ	23 FEP MA = 24.5 (5.6) Male: 47.82% Yrs Ed: 12.3 (1.2) Months Ill: (36.4) (34.5) Clor eq: 539.7 (296.9) WAIS VT: 21.8 (5.9)	23 HC MA = 26.8 (4.2) Male: 57.14% Yrs Ed: 12.7 (0.6) WAIS VT: 22.2 (5.2)	ToM	Moving Shapes Paradigm (Visual ToM, animated shapes task) (Abell et al., 2000)	Intentionality score (ToM vocabulary) significantly correlated with positive subscale of the PANSS ($r = 0.57$, $p < 0.01$); no other significant correlations	Auditory Verbal Learning and Memory; Reasoning; cognitive flexibility	• Significant correlations between neurocognitive indices and ToM indices (r's ranging from 0.52–0.55, $p < 0.05$) • Series ANCOVAs with ToM indices as DV (appropriateness, intentionality) and different neurocognitive indices as covariates • FEP vs. HC, intentionality scores: n.s. when the following neurocognitive indices were used as covariates: LPS3 (reasoning) ($F(2,42) = 1.43$, $p = 0.25$); auditory verbal learning and memory test ($F(4,40) = 1.11$, $p = 0.34$). • HC > FEP, Intentionality was not affected by WAIS-R score ($F(4,40) = 3.58$, $p < 0.05$) • FEP vs. HC, appropriateness scores: HC > FEP, appropriateness scores were not by for any neuropsychological indices ($p < 0.05$)
Langdon, Connors, et al. (2014)	FEP: <2 years of first presentation to tx. Dx: SCZ ($n = 21$), SCZA ($n = 1$), PNOS ($n = 1$)	23 FEP, MA = 20.91 (1.83) Male: 95.65% Yrs Ed: 11.43 (2.02) Wks Ill: 50.74 (29.50) Age of onset:	19 HC MA = 20.79 (1.81) Male: 89.47% Yrs Ed: 12.82 (1.94) IQ: 103.42 (9.32)	ToM	Non-verbal picture sequencing task (Visual ToM) (Langdon & Coltheart, 1999) Joke Appreciation Task (Visual ToM) (Happé, Brownell, & Winner, 1999)	ToM composite significantly correlated with negative symptoms ($r = -0.50$, $p < 0.05$).	Composite score of neurocognition (with ToM composite only)	Picture sequencing task • FEP vs. HC: HC > FEP, $t(40) = 3.81$, $p < 0.01$ Joke Appreciation

		19.91 (1.95) Clor eq: IQ: 96.65 (8.41)			Story comprehension test (Verbal ToM) (Happé et al., 1999)			• FEP vs. HC: HC > FEP, $t(40) = 5.23$, $p < 0.01$ Story Comprehension Task • FEP vs. HC: HC > FEP, $t(40) = 5.17$, $p < 0.01$ *All total scores, no subscales* ToM Composite • ToM indices were significantly correlated and thus collapsed into a composite score • FEP vs, HC: HC > FEP; ANCOVA with ToM composite, controlling for effect of neurocognitive composite score $F(1,38) = 5.60$, $p = 0.02$
Langdon, Still, et al. (2014)		Same sample as Langdon, Connors, et al., 2014: 23 FEP, MA = 20.91 (1.83) Male: 95.65% Yrs Ed: 11.43 (2.02) Wks Ill: 50.74 (29.50) Age of onset: 19.91 (1.95) Clor eq: IQ: 96.65 (8.41)	Same sample as Langdon, Connors, et al., 2014: 19 HC MA = 20.79 (1.81) Male: 89.47% Yrs Ed: 12.82 (1.94) IQ: 103.42 (9.32)	ToM	Non-verbal picture sequencing task (Visual ToM) (Langdon & Coltheart, 1999) Joke Appreciation Task (Visual ToM) (Happé, Brownell, & Winner, 1999) Story comprehension test (Verbal ToM) (Happé et al., 1999)	NR	Estimate of IQ using reading test (NART)	Picture sequencing task: • FEP vs. HC: HC > FEP; post-hoc comparisons HC > FEP on ToM vs. control sequences ($F(1,40) = 7.159$, $p = 0.011$) Joke appreciation task: • FEP vs. HC: HC > FEP; post-hoc comparisons, FEP significantly worse on ToM jokes than control jokes ($F(1,40) = 12.450$, $p = 0.001$) Story Comprehension Task • FEP vs. HC: n.s. differences regarding type of story, ToM and control. ($F(1,39) = 0.535$, $p = 0.469$) Estimated IQ: • Estimated IQ not significantly related to task performance for picture sequencing task and joke appreciation task ($ps > 0.675$) • Estimated IQ, trend level significance w/story comprehension task ($F(1,39) = 3.907$, $p = 0.055$)
Mazza et al. (2012)	FEP: <3 months of initial dx (e.g. first time of hospitalization or presentation for services). Recruited within a month of presentation. Dx: SCZ, SCZA, SZF	49 FEP, MA = 26.8 (6.4) Male: 67.35% Yrs Ed: 10.8 (2.8) Yrs Ill: 1.01 (0.5) Clor eq: NR IQ: NR	386 HC (undergraduates); matched for age and education, demographics NR	ToM	Happé's Strange Stories (Verbal ToM) (Happé, 1994) False Belief Task (Verbal ToM) (Mazza et al., 2012)	Happé's Strange Stories significantly correlated with BPRS Negative symptom subscale ($r = 0.38$, $p < 0.000$)	None	Happé's Strange Stories • FEP vs. HC: HC > FEP, post-hoc Bonferroni correction, undergraduate HC sample (mean differences $= -2.40$, $p < 0.001$); community HC sample (mean differences $= -1.76$, $p < 0.001$, effect size, $d = 1.02$) • FEP vs. SCZ: n.s. differences False Belief Task • FEP vs. HC, first-order: HC > FEP, post-hoc Bonferroni correction, undergraduate HC
	SCZ: chronic SCZ Dx: SCZ	178 SCZ MA = 34.5 (8.4) Male: 69.66% Yrs Ed: 11.54	96 HC (community) matched for age and education					

(continued on next page)

Appendix B (*continued*)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
		(2.7) Yrs Ill: 10.9 (2.9) Clor eq: 210.3 mg/day (143.67) IQ: NR Avg. hosp: NR						sample (mean differences = −0.891, $p < 0.001$); community HC sample (mean differences = −0.875, $p < 0.001$) • FEP vs. SCZ, first-order: n.s. differences • FEP vs. HC, second-order: HC > FEP, post-hoc Bonferroni correction, undergraduate HC sample (mean differences = −0.891, $p < 0.001$); community HC sample (mean differences = −0.875, $p < 0.001$) • FEP vs. SCZ, second-order: n.s. differences
Mazza et al. (2013)	FEP: within 3 months of initial dx (e.g. first time of hospitalization or presentation for services) Dx: SCZ ($n = 4$), SCZA ($n = 4$), PNOS ($n = 3$), delirium ($n = 1$)	12 FEP, MA = 27.92 Male: 66.7% Yrs Ed: 13.5 (2.3) Yrs Ill: NR Clor eq: NR IQ: NR	12 HC MA = 27.25 (9.3) Male: 41.67% Yrs Ed: 15 (2.3) IQ: NR	EP ToM	The Eyes Task (Visual ToM, human stimuli) (Baron-Cohen et al., 2001)	NR	None	FEP vs. HC: HC > FEP, HC performed significantly better on the Eyes Task ($t(1,24) = -4.26$, $p < 0.001$)
Montreuil et al. (2010)	FEP: age 14–30 Inpatients and outpatients; FE patients; no medications >1 month Dx: SZ spectrum, affective psychosis, PNOS	45 FEP, subgroups: Poor outcome (n = 27) MA = 23.5 (3.7) Male: 70.37% Yrs Ed: 11.6 (2.8) DUP: 38.3 (44.3) weeks Clor eq: NR IQ: NR Good outcome ($n = 18$) MA = 23.9 (3.0) Male: 66.67% Yrs Ed: 12.2 (2.5) DUP: 93.7 (148.2) weeks Clor eq: NR IQ: NR	26 HC MA = 24.7 (3.6) Male: 53.87 Yrs Ed: 14.4 (1.7) IQ: NR	ToM SP	Hinting Task (Verbal ToM) (Corcoran et al., 1995)	No significant associations	Parental SES, verbal memory, working memory	• FEP vs. HC: HC > FEP, both FEP outcome subgroups. Fisher's LSD comparisons, bonferroni correction set at $p < 0.01$. (p's range from 0.001–0.007) • ANCOVA controlling for verbal memory and working memory; results remained unchanged (no statistics reported)
Thompson et al. (2012)	FEP: 1 week of daily psychotic symptoms (Comprehensive Assessment of At Risk Mental States; Yung 2005); <6 months tx for a psychotic disorder	40 FEP, outpatients MA = 20.5 (2.5) Male: 62.5% Yrs Ed: 12.5	30 HC MA = 19.3 (2.9) Male: 41.9% Yrs Ed: 13.2 IQ: 104.9 (12.5) WASI	EP (higher-level, lower-level, face and voice) ToM	Hinting Task (Verbal ToM) (Corcoran et al., 1995) Interpretation of Visual Jokes (Visual ToM) (Corcoran et al., 1997)	NR	Age, sex, IQ estimated based on Reading Test (NART)	Hinting Task • FEP vs. HC: HC > FEP, Post-hoc Tukey's tests adjusted for multiple testing; FEP performed significantly worse than

		(1.6) Yrs Ill: NR Clor eq: NR Meds: 30/40 on Antipsychotics Estimated IQ: 106.4 (10.6) (based on NART)		SP				controls, even after controlling for age, gender, and IQ ($p < 0.01$) • Effect size = 0.93 Interpretation of Visual Jokes • FEP vs. HC: Post-hoc Tukey's tests adjusted for multiple testing; HC > FEP after controlling for age, gender, and IQ in: overall score and physical (control) score ($p < 0.01$) • FEP vs. HC: n.s. in ToM specific (mentalizing) scale not significant. • Total score effect size, 0.64
Ventura et al. (2015)	FEP: recent-onset psychosis < 2 years of study entry (recency of episode; mean = 7.1 months). 87% current FEP Dx: SCZ ($n = 50$), SCZA ($n = 7$), SZF ($n = 20$)	77 FEP, MA = 21.47 (3.76) Male: 78% Yrs Ed: 12.36 (1.75) Yrs Ill: 7.1 months Clor eq: NR Cognitive composite: 29.16 (14.42)	21 HC MA = 22.68 (2.34) Male: 63% Yrs Ed: 14.11 (1.91) Cognitive composite: 45.73 (8.66)	ToM	Silent animations task (Visual ToM) (Abell et al., 2000; Castelli et al., 2000)	Both ToM Appropriateness/Intentionality significantly correlated with SANS negative symptoms at baseline and retest (r's range − 0.24 through −0.50), $p < 0.05$). Positive symptoms significantly associated at follow up only (r's range − 0.29 to −0.38, $p < 0.05$)	Parental education, ethnicity	• FEP vs. HC, ToM Videos: HC > FEP, Intentionality score ($F(1,88) = 15.9$, $p < 0.00$); appropriateness score ($F(1,88) = 12.7$, $p < 0.00$) • FEP vs. HC, Random Videos (control): n.s. differences between groups • FEP vs. HC, Goal-Directed Videos (control): HC > FEP on intentionality score ($F(1,89) = 4.13$, $p = 0.04$); n.s. difference on appropriateness score • Symptom remitted subgroup comparisons: Symptom remitted patients (<3 on all BPRS positive symptoms) at 6 months ($n = 31$) compared to HCs • FEP vs. HC, remitted subgroup: still evidenced a deficit on ToM intentionality ($F(1,46) = 4.82$, $p = 0.03$, and ToM appropriateness ($F(1,46) = 4.23$, $p = 0.04$)
Vohs et al. (2014)	FEP: age 18–65; <5 years of a documented initial dx of a SZ spectrum disorder; no medication changes or hospitalizations within 30 days Dx: SCZ, SCZA, PNOS	26 FEP, MA = 23.81 (3.63) Male: 81% Yrs Ed: 13 (1.87) Yrs Ill: NR Clor eq: 307.2 (60.2) IQ: NR	Used non-psychiatric control group from substance use disorders program	EP ToM	Hinting Task (Verbal ToM) (Corcoran et al., 1995) The Eyes Task (Visual ToM, human stimuli) (Baron-Cohen et al., 2001)	Eyes and Hinting Task, respectively: not significantly associated with positive = sx ($r = -0.03$; $r = -0.30$). Associated with negative sx ($r = -0.64$, $p < 0.001$; $r = 0.60$, $p < 0.01$) and disorganized sx ($r = -0.68$, $p < 0.001$; $r = -0.64$, $p < 0.001$)	Age	FEP vs. SCZ: n.s. differences between FEP and SCZ for both tasks.
	SCZ: age 18–65; >5 years documented initial dx of a SCZ spectrum disorder; no medication changes or hospitalization within 30 days Dx: SCZ, SCZA	72 SCZ MA = 51.01 (6.94) Male: 88% Yrs Ed: 12.72 (1.81) Yrs Ill: NR Clor eq: 420.5 (93.4) IQ: NR Avg. hosp: NR						

ANCOVA = Analysis of Covariance; AS = Attributional Style; Clor eq = Chlorpromazine equivalent (mg/d); DART = Dutch Adult Reading Test; DD = Delusional Disorder; DUP = Duration of Untreated Psychosis; Dx = Diagnosis; Edu = Education; EP = Emotion Processing; FE: First Episode; FEP = First Episode Psychosis; HC = Healthy Control; HS = High school; MA = Mean Age; NART = National Adult Reading Test; NR = Not Reported; n.s. = non-significant; PANSS: Positive and Negative Syndrome Scale; PNOS = Psychosis NOS; SCZ = Schizophrenia; SCZA = Schizoaffective; SP = Social Perception; SZF = Schizophreniform; Sx = symptoms; TASIT = The Awareness of Social Inference Test; ToM = Theory of Mind; Tx = Treatment; WAIS-R = Wechsler Adult Intelligence Scales-Revised; Yrs Ed = Years of Education; Yrs Ill = Number of Years Ill.

Appendix C. Social Perception (SP) in First Episode Psychosis (FEP)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Covariates	Main findings
Achim et al. (2012)	FEP: age 18–35; in early stages of a psychotic disorder Dx: SCZ (n = 23), SCZA (n = 2), DD (n = 4), PNOS (n = 2)	31 FEP, MA = 24.9 (4.5) Male: 83.87% Edu category: Hollingshead's categories; 4.0 (1.1) Yrs Ill: NR Duration of tx: 20.9 months Clor eq: NR Estimated IQ: 100.4 (15.1)	31 HC, MA = 25.2 (4.2) Male: 83.87% Yrs Ed: Hollingshead categories; 3.3 (1.2) IQ: 101.8 (10.5)	ToM EP SP	Social Knowledge Task (Achim et al., 2012)	NR	Edu	• FEP vs. HC: no significant differences between groups ($t(60) = 1.49$, $p = 0.14$
Addington et al. (2006b)	FEP: currently experiencing FEP; <3 months of treatment	Same sample as Addington et al., 2006a: 50 FEP, MA = 25.6 (8.0) Male: 60% Completed HS: 66% Yrs Ill: NR Clor eq: 380 mg/day IQ: NR	Same sample as Addington et al., 2006a: 55 HC, MA = 21.2 (6.1) Male: 60% Completed HS: 72.2 IQ: NR	SP	SCRT (Corrigan, 1997) SFRT (Corrigan & Green, 1993)	Patient group (FEP and SCZ combined); significantly associated with positive and negative sx at both baseline and retest (r values ranged from 0.27, $p < 0.01$, to 0.38 $p < 0.0005$).	None	SCRT/SFRT • FEP vs. SCZ: n.s. differences on either task • FEP/SCZ vs. HC: patient groups collapsed because of similar performance (n = 103). HC > FEP/SCZ, F values ranged from 9.45 to 18.60, $p < 0.0001$ for each model (df and specific statistics not reported) • Significant associations between FEDT and cognition at all time points
	SCZ: dx >3 years; >2 admissions to the hospital	Same sample as Addington et al., 2006a: 53 SCZ MA = 35.5 (7.2) Male: 72% Completed HS: 71.7% Yrs Ill: NR Clor eq: 715 mg/day IQ: NR Avg. hosp: 5						
Bertrand et al. (2007)	FEP: patients with a FE. Dx: SCZ (n = 21), SCZA (n = 8), PNOS (n = 6), and SZF (n = 1)	36 FEP, MA = 22.78 (3.37) Male: 64% Yrs Ed: NR Yrs Ill: NR Days since beginning tx: mean = 143; range 0–529 days DUP (weeks): 62.24 (86.95) range 1–310 Clor eq: NR IQ: 92.31 (15.94)	25 HC MA = 24.19 (3.55) Male: 52% Yrs Ed: NR IQ: 106.31 (10.90)	ToM SP	Four Factor Tests of Social Intelligence (O'Sullivan & Guilford, 1976)	No significant associations	IQ	• FEP vs. HC: HC > FEP, ANCOVA for Four Factor Test of Social Intelligence; controlling for IQ, significant effect of group ($F(1,57) = 11.08$, $p = 0.002$) • Effect held when looking at schizophrenia diagnosis participants only ($F(1,43) = 19.04$, $p < 0.0001$) Regarding subscales, separate t-tests were applied to each, significant group differences across (all $ps < 0.001$)

Green et al. (2012)	FEP: age 18–45; AAO within 2 years of study participation; most recovering from FE (e.g. approx 2–3 months after hosp, after med level was stabilized) Dx: SCZ ($n = 46$), SCZA ($n = 10$), SZF ($n = 25$) SCZ: FE >5 years before testing Dx: SCZ ($n = 48$), SCZA ($n = 5$)	81 FEP MA = 22.02 (4.18) Male: 75% Yrs Ed: 12.50 (1.96) Yrs Ill: NR Clor eq: NR IQ: NR 53 SCZ MA = 34.77 (7.89) Male: 66% Yrs Ed: 13.96 (1.64) Yrs Ill: NR Clor eq: NR IQ: NR Avg. hosp: NR	46 HC MA = 22.20 (3.51) Male: 63% Yrs Ed: 13.86 (1.97) IQ: NR 47 SCZ HC MA = 33.02 (5.32) Male: 72% Yrs Ed: 14.45 (1.69) IQ: NR	EP SP	RAD (Fiske, 1992)	Negative sx were significantly correlated with the RAD ($r = -0.30$, $P < 0.05$)	Parental education	• FEP vs. HC: HC > FEP, Cohen's $d = 1.02$ ($p < 0.01$) • FEP vs. SCZ: group (clinical, control) by phase (CHR, FEP, SCZ) indicates change over the course of illness, significant for RAD ($F(2282) = 3.364$ $p = 0.028$). • Though significant, effect was small (eta square = 0.025) • Inspection, this interaction is due to decreased performance in FEP
Montreuil et al. (2010)	FEP: age 14–30 inpatients/outpatients; FE patients; no medications >1 month Dx: SCZ spectrum, affective psychosis, PNOS	45 FEP, subgroups: Poor outcome ($n = 27$) MA = 23.5 (3.7) Male: 70.37% Yrs Ed: 11.6 (2.8) DUP: 38.3 (44.3) weeks Clor eq: NR IQ: NR Good outcome ($n = 18$) MA = 23.9 (3.0) Male: 66.67% Yrs Ed: 12.2 (2.5) DUP: 93.7 (148.2) weeks Clor eq: NR IQ: NR	26 HC MA = 24.7 (3.6) Male: 53.87 Yrs Ed: 14.4 (1.7) IQ: NR	ToM SP	Four Factor Tests of Social Intelligence (O′Sullivan & Guilford, 1976)	No significant associations	Parental SES, verbal memory, working memory	• FEP vs. HC: HC > FEP, both FEP outcome subgroups. Fisher's LSD comparisons, bonferroni correction set at $p < 0.01$. (p's < 0.001), except missing cartoon subtest (n.s.) • ANCOVA controlling for verbal memory and working memory; results remained unchanged (no statistics reported)

Confounds include variables that are controlled for in subsequent analyses. ANCOVA = Analysis of Covariance; AS = Attributional Style; CHR = Clinical High Risk; Clor eq = Chlorpromazine equivalent (mg/d); DD = Delusional Disorder; DUP = Duration of Untreated Psychosis; Dx = Diagnosis; EP = Emotion Processing; FE: First Episode; FEP = First Episode Psychosis; HC = Healthy Control; HS = High school; MA = Mean Age; NR = Not Reported; n.s. = non-significant; PANSS: Positive and Negative Syndrome Scale; PNOS = Psychosis NOS; RAD = Relationships Across Domains; SCRT = Social Cue Recognition Test; SCZ = Schizophrenia; SCZA = Schizoaffective; SES = Socioeconomic Status; SFRT = Situational Features Recognition Test; SP = Social Perception; SZF = Schizophreniform; Sx = symptoms; ToM = Theory of Mind; Tx = Treatment; WAIS-R = Wechsler Adult Intelligence Scales-Revised; Yrs Ed = Years of Education; Yrs Ill = Number of Years Ill.

Appendix D. Attributional Style (AS) in First Episode Psychosis (FEP)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Covariates	Main findings
Achim et al. (2016)	FEP: recent-onset schizophrenia spectrum disorder	29 FEP MA = 25.4 (4.8) Male: 12%	41 HC MA = 25.1 (4.3) Male: 37%	AS	Brief-IPSAQ (Kinderman & Bentall, 1997)	No significant correlations between externalizing bias/personalizing biases and PANSS delusions (P1) and suspiciousness (P6) items. No other symptom correlations reported.	None	• FEP vs. HC, externalizing bias: n.s. difference ($t(68) = -0.8$, $p = 0.42$) • FEP vs. HC, personalizing bias: n.s. difference (not reported) • FEP with delusions vs. without: greater personalizing in the delusions group (PANSS score > 2) ($t(39) = 2.78$, $p = 0.008$)
An et al. (2010)	FEP: from outpatient ($n = 15$) and inpatient ($n = 5$); age 15–35 at onset of psychosis Dx: SCZ; excluded with >1 admission to psychiatric hospital or >12 months on an antipsychotic medication	20 FEP, MA = 21.3 (5.0) Male: 40% Yrs Ed: 12.6 (2.5) Months Ill: 11.7 (11.5) Weeks of meds: 12 (13.1) Clor eq: 589.5 (369.0) IQ: NR	39 HC MA = 19.7 (3.5) Male: 41% Yrs Ed: 12.4 (2.0) IQ: NR	AS	AIHQ (Korean translation) (Combs et al., 2007)	No significant correlations between positive, negative, or general subscales of the PANSS. Hostility bias significantly correlated with the paranoia scale of the PANSS ($r = 0.57$, $p = 0.001$). Trend level association with suspiciousness persecution PANSS item ($r = 0.36$, $p = 0.12$).	None	• FEP vs. HC: Post hoc testing with Bonferroni correction showed that FEP evidenced a greater hostility bias than HC ($p = 0.037$). n.s. differences between FEP and HC on blame or aggression biases
Fornells-Ambrojo and Garety (2009)	FEP: age 16–45; largely inpatients. Early psychosis participants; first presentation of psychosis <5 years; patients in their second ($n = 12$) episode of psychosis included; required a PANSS score of 4 (Moderate) on suspiciousness item. Dx: SCZ, SCZA, SZF	20 FEP, MA = 27.2 (7.9) Male: 90% Yrs Ed: 95% secondary; 5% university Months Ill: 7.3 (6.3) Clor eq: NR IQ: NR	32 HC MA = 26.7 (5.3) Male: 81% Yrs Ed: 56% secondary; 44% university IQ: NR	AS	ARAT (Fornells-Ambrojo and Garety, 2009)	NR	None	• FEP vs. HC, self-serving bias: n.s. difference in tendency to self-blame instead of taking credit for success (Scheffe = 0.87, $p = 0.2$) • FEP vs. HC, other-person bias: FEP are more likely to blame other people for negative events rather than themselves compared to HC (Scheffe = 0.19; $p = 0.014$)
So et al. (2015)	FEP; all outpatients. Excluded for comorbid substance use d/o, mood d/o, organic mental d/o, intellectual disability	70 FEP MA = 20.0 (3.3) Male: 47% Yrs ed: 10.8 (2.3) Months of tx: 20.2 (range: 0–48)	642 HC MA = 21.0 (4.1) Male: 42% Yrs ed: 11.0 (2.0)	AS	IPSAQ (Kinderman & Bentall, 1996)	Self-serving bias significantly associated with number of delusional beliefs ($t = 6.7$, $p < 0.001$) delusional conviction ($t = 7.118$, $p < 0.001$), distress ($t = 7.3$, $p < 0.001$), and preoccupation ($t = 6.4$, $p < 0.001$) Personalizing bias significantly associated with number of delusional beliefs ($t = 3.9$, $p < 0.001$) delusional conviction ($t = 7.4$, $p < 0.001$), distress ($t = 7.3$, $p < 0.001$), and preoccupation ($t = 7.7$, $p < 0.001$)	None	• FEP vs. HC, self-serving bias: bonferroni adjusted pairwise comparison, FEP > HC (MD = 9.60, SE = 2.28, $p < 0.001$) • FEP vs. HC, personalizing bias: n.s. after adjusting for multiple comparisons (MD = 0.06, SE = 0.03, $p = 0.044$; using a p value of >0.025)

Confounds include variables that are controlled for in subsequent analyses. AIHQ = Ambiguous Intentions Hostility Questionnaire; ARAT = Achievement and Relationships Attributions Task; AS = Attributional Style; Clor eq = Chlorpromazine equivalent (mg/d); DD = Delusional Disorder; Dx = Diagnosis; EP = Emotion Processing; FE: FEIT = Face Emotion Identification Task; FEDT = Face Emotion Discrimination Task; First Episode; FEP = First Episode Psychosis; HC = Healthy Control; HS = High school; MA = Mean Age; NR = Not Reported; n.s. = non-significant; PANSS: Positive and Negative Syndrome Scale; PNOS = Psychosis NOS; SCZ = Schizophrenia; SCZA = Schizoaffective; SP = Social Perception; SZF = Schizophreniform; Sx = symptoms; ToM = Theory of Mind; Tx = Treatment; WAIS-R = Wechsler Adult Intelligence Scales-Revised; Yrs Ed = Years of Education; Yrs Ill = Number of Years Ill.

Appendix E. Longitudinal studies of social cognition in First Episode Psychosis (FEP)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
Addington et al. (2006a)	FEP: currently experiencing FEP; <3 months of treatment SCZ: dx >3 years; >2 admissions to the hospital	50 FEP, MA = 25.6 (8.0) Male: 60% Completed HS: 66% Yrs Ill: NR Clor eq: 380 mg/day Cog: NR 53 SCZ MA = 35.5 (7.2) Male: 72% Completed HS: 71.7% Yrs Ill: NR Clor eq: 715 mg/day IQ: NR Avg. hosp: 5	55 HC, MA = 21.2 (6.1) Male: 60 Completed HS: 72.2 IQ: NR	EP (lower-level)	FEIT FEDT (Kerr & Neale, 1993)	NR	None	FEIT/FEDT composite score; 1 year • SCZ/FEP: no significant changes in FEDT/FEIT composite score, combined patient groups ($t(102) = -1.50$, p = n.s.) • HC: significant increase in FEIT over 1 year (FEIT: $t = -3.27$, $p < 0.002$)
Addington et al. (2006b)	FEP: currently experiencing FEP; <3 months of treatment SCZ: dx >3 years; >2 admissions to the hospital	Same sample as Addington et al., 2006a: 50 FEP, MA = 25.6 (8.0) Male: 60% Completed HS: 66% Yrs Ill: NR Clor eq: 380 mg/day Cog: NR Same sample as Addington et al., 2006a: 53 SCZ MA = 35.5 (7.2) Male: 72% Completed HS: 71.7% Yrs Ill: NR Clor eq: 715 IQ: NR Avg. hosp: 5	Same sample as Addington et al., 2006a: 55 HC, MA = 21.2 (6.1) Male: 60 Completed HS: 72.2 IQ: NR	SP	SCRT (Corrigan, 1997) SFRT (Corrigan & Green, 1993)	Patient group (FEP and SCZ combined); significantly associated with positive and negative sx at both baseline and retest (r values ranged from 0.27, $p < 0.01$, to 0.38 $p < 0.0005$).	None	1 year longitudinal study • SCZ/FEP: no significant changes in SCRT or SFRT (p is n.s.) • HC: significant increase over time for SCRT ($t = -3.13$, $p < 0.01$) and SFRT ($t = -4.04$, $p < 0.0005$)
Hill et al. (2008)	FEP: Recently experienced an acute episode of psychosis that required tx; must be ill <5 years Dx: SCZ, SZF, SCZA	219 FEP, MA = 24.45 Male: 71.7% Yrs Ed: NR Yrs Ill: 13.53 months (17.91) Clor eq: NR WRAT-III: 94.68 (16.92)	None	EP	Penn Emotion Discrimination Test (Erwin et al., 1992)	NR	None	12 week longitudinal study Evidence that EP ability was stable over time, ICCs of 0.68
Horan et al. (2012)	FEP: age 18–45; AAO within 2 years of study participation; most recovering from FE (e.g. approx. 2–3 months after hosp, after med level was stabilized)	55 FEP, MA = 22.3 (4.3) Male: 75% Yrs Ed: 12.7 (2.1)	None	SP ToM	MSCEIT (Mayer et al., 2003) TASIT (part III) (McDonald et al., 2006)	Symptom analyses used a collapsed SC composite (TASIT III, MSCEIT, and RAD):	None	1 year longitudinal study MSCEIT

(continued on next page)

Appendix E (*continued*)

Study	Group criteria	Patient group	Control group	Domains	Measures	Symptom correlations	Confounds	Main findings
		Yrs Ill: NR Clor eq: NR WAIS-R: NR Hosp #: 1.2 (0.8) months ill: 8.5 (6.4) DUP: 6.9 (6.2)			RAD (Fiske, 1992)	• Significantly correlated with total sx (BL: $r = -0.36$, $P < 0.05$; FU: $r = -0.41$, $P < 0.01$). • Significantly correlated with positive sx (BL: $r = -0.33$, $P < 0.05$; FU: $r = -0.31$, $P < 0.05$) Significantly correlated with negative sx at FU only ($r = -0.33$; $P < 0.05$)		• $t = 1.11$, p is n.s. • $d = 0.16$ • $r = 0.87$, $p < 0.001$ TASIT • $t = 2.50$, $p < 0.05$ • $d = 0.36$ • $r = 0.71$, $p < 0.001$ RAD • $t = 2.43$, $p < 0.05$ • $d = 0.35$ • $r = 0.74$, $p < 0.001$ Summary • Small to medium effect sizes across tasks. Paired samples t-tests showed significant improvement for each. MSCEIT did not show a significant improvement at follow-up
Sullivan et al. (2014)	FEP: Recent onset psychosis, entering secondary care for the first time with psychotic sx Dx: PNOS (34%), SCZ (11%), drug-induced psychosis (10%), MDD w/psy features (6%), BPD with mania (6%), SCZA (3%), DD (2%)	34 FEP, MA = 24.4 (7.1) Male: 67.7% Yrs Ed: dichotomized variable; 42.4% higher education Yrs Ill: NR DUP: 23.28 months Clor eq: IQ: 105.2 (10.0) (NART CONVERSION)	None	ToM	Hinting Task (Verbal ToM) (Corcoran et al., 1995) Interpretation of Visual Jokes Task (Corcoran et al., 1997)	NR	Symptoms	1 year longitudinal study • Evidence that ToM ability was stable over time, ICCs as follows: Hinting task (0.46), Visual Cartoon ToM jokes (0.38), physical jokes (0.49) (no p values provided)
Ventura et al. (2015)	FEP: recent-onset schizophrenia; onset of psychosis <2 years of study entry (recency of episode; mean $n = 7.1$ months). 87% were currently experiencing FEP Dx: SCZ ($n = 50$), SCZA ($n = 7$), SZF ($n = 20$)	31 FEP, Demographics not reported for subsample, but original sample ($n = 77$)	None	ToM	Silent animations task (Visual ToM) (Abell et al., 2000; Castelli et al., 2000)	Both ToM appropriateness/intentionality significantly correlated with SANS negative symptoms at baseline and retest (r's range − 0.24 through −0.50), $p < 0.05$). Positive symptoms Significantly associated at follow up only (r's range − 0.29 to −0.38, $p < 0.05$)	Parental education, ethnicity	6 month longitudinal study • ToM scenes significantly correlated BL and FU; intentionality ($n = 48$, $r = 0.57$, $p < 0.01$), appropriateness ($n = 48$, $r = 0.41$, $p < 0.01$) • ToM scenes change between BL and FU: trend level decrease in intentionality ($t(47) = -1.95$, $p = 0.06$), appropriateness did not change ($t(47) = -0.41$, $p = 0.69$)

Confounds include variables that are controlled for in subsequent analyses. AS = Attributional Style; BL = Baseline; Clor eq = Chlorpromazine equivalent (mg/d); DD = Delusional Disorder; DUP = Duration of Untreated Psychosis; Dx = Diagnosis; EP = Emotion Processing; FE: First Episode; FEP = First Episode Psychosis; FU = Follow up; HC = Healthy Control; HS = High school; MA = Mean Age; MDD = Major Depressive Disorder; MSCEIT = Mayer-Salovey-Caruso Emotional Intelligence Test; NART = National Adult Reading Test; NR = Not Reported; n.s. = non-significant; PANSS: Positive and Negative Syndrome Scale; PNOS = Psychosis NOS; RAD = Relationships Across Domains; SCRT = Social Cue Recognition Test; SCZ = Schizophrenia; SCZA = Schizoaffective; SFRT = Situational Features Recognition Test; SP = Social Perception; SZF = Schizophreniform; Sx = symptoms; TASIT = The Awareness of Social Inference Test; ToM = Theory of Mind; Tx = Treatment; WAIS-R = Wechsler Adult Intelligence Scales-Revised; Yrs Ed = Years of Education; Yrs Ill = Number of Years Ill.

Appendix F. References in appendices

Abell, F., Happe, F., Frith, U. (2000). Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. Cognitive Development, 15, 1–16.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Aspberger Syndrome or high-functioning Autism. Journal of Child Psychology and Psychiatry, 42, 241–251.

Bediou, B., Krolak-Salmon, P., Saoud, M., Henaff, M., Burt, M., Dalery, J., & D'Amato,

T. (2005). Facial expression and sex recognition in schizophrenia and depression. Canadian Journal of Psychiatry, 50, 525.

Blair, R. J., & Cipolotti, L. (2000). Impaired social response reversal: A case of 'acquired sociopathy'. Brain, 123, 1122–1141.

Calder, A. J., Rowland D., Young A. W., Nimmo-Smith I., Keane J., & Perrett D. I. (2000). Caricaturing facial expressions. Cognition, 76, 105–146.

Castelli F., Happé, F., Frith, U., Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. NeuroImage, 12, 314–325.

Corcoran, R., Mercer, G., Frith, C.D. (1995). Schizophrenia, symptomatology and social inference: Investigating "theory of mind" in people with schizophrenia. Schizophrenia Research, 17, 5–13.

Corrigan, P. (1997). The social perceptual deficits of schizophrenia. Psychiatry, 60, 309-326.

Corrigan, P., & Green M. F. (1993) The situational feature recognition test: A measure of schema comprehension for schizophrenia. International Journal of Methods in Psychiatric Research, 3, 29–35.

Ekman, P. (1989). The argument and evidence about universal facial expressions of emotion. In K. R. Scherer & P. Ekman (Eds.), Approaches To Emotion. New York: Oxford University Press.

Ekman, P., Friesen, W.V. 1976. Pictures of Facial Affect. Consulting Psychologists Press, Palo Alto.

Erwin, R. J., Gur, R. C., Gur, R. E., Skolnick, B., Mawhinney-Hee, M., & Smailis, J.

(1992). Facial emotion discrimination: I. Task construction and behavioral findings in normal subjects. Psychiatry Research, 42, 231–240.

Feinberg, T.E., Rifkin, A., Schaffer, C., Walker, E. (1986). Facial discrimination and emotional recognition in schizophrenia and affective disorders. Archives of General Psychiatry, 43, 276–279.

Fiske A. P. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. Psychological Review, 99, 689–723.

Fornells-Ambrojo, M., & Garety, P. A. (2009). Attributional biases in paranoia: The development and validation of the Achievement and Relationships Attributions Task (ARAT). Cognitive Neuropsychiatry, 14, 87–109.

Gur, R. C., Ragland, J. D., Moberg, P. J., Turner, T. H., Bilker, W. B., Kohler, C., … & Gur, R. E. (2001). Computerized neurocognitive scanning: I. Methodology and validation in healthy people. Neuropsychopharmacology, 25, 766–776.

Gur, R. C., Sara, R., Hagendoorn, M., Marom, O., Hughett, P., Macy, L., … & Gur, R. E. (2002). A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. Journal of Neuroscience Methods, 115, 137–143.

Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. Journal of Autism and Developmental Disorders, 24, 129–154.

Happé, F., Brownell, H., & Winner, E. (1999). Acquired theory of mind impairments following stroke. Cognition, 70, 211–240.

Inoue, Y., Tonooka, Y., Yamada, K., & Kanba, S. (2004). Deficiency of theory of mind in patients with remitted mood disorder. Journal of Affective Disorders, 82, 403–409.

Kerr, S. L., & Neale, J. M. (1993). Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance?. Journal of Abnormal Psychology, 102, 312.

Kucharska-Pietura, K., Phillips, M. L., Gernand, W., & David, A. S. (2003). Perception of emotions from faces and voices following unilateral brain damage. Neuropsychologia, 41, 1082–1090.

Langdon, R., & Coltheart, M. (1999). Mentalising, schizotypy, and schizophrenia. Cognition, 71, 43–71.

Matsumoto, D., & Ekman, P. (1995). Japanese and Caucasian facial expressions of emotion (JACFEE) and neutral faces (JACNeuF).

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. Emotion, 3, 97.

McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., Neulinger, K. (2006).

Reliability and validity of The Awareness of Social Inference Test (TASIT): A clinical test of social perception. Disability Rehabilitation, 28, 1529–1542.

Nowicki Jr., S., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. Journal of Nonverbal behavior, 18, 9–35.

O'Sullivan, M., & Guilford, J. P. (Eds.) (1976). Four Factor Tests of Social Intelligence (Behavioral Cognition): Manual of Instructions and Interpretations. Sheridan Psychological Services, Inc., Orange, CA.

Schneider F., Gur, R. C., Koch, K., Backes, V., Amunts, K., Shah, N. J., … & Habel, U. (2006). Impairment in the specificity of emotion processing in schizophrenia. American Journal of Psychiatry, 163, 442–447.

Wechsler, D. (1999). Wechsler abbreviated scale of intelligence. Psychological Corporation.

Williams, L. M., Palmer, D., Liddell, B. J., Song, L., & Gordon, E. (2006). The 'when' and 'where' of perceiving signals of threat versus non-threat. Neuroimage, 31(1), 458–467.

Williams, L. M., Kemp, A. H., Felmingham, K., Liddell, B. J., Palmer, D. M., & Bryant, R. A. (2007). Neural biases to covert and overt signals of fear: dissociation by trait anxiety and depression. Journal of Cognitive Neuroscience, 19(10), 1595–1608.

References

Achim, A. M., Ouellet, R., Roy, M. A., & Jackson, P. L. (2012). Mentalizing in first-episode psychosis. *Psychiatry Research*, *196*(2), 207–213.

Achim, A. M., Sutliff, S., Samson, C., Montreuil, T. C., & Lecomte, T. (2016). Attribution bias and social anxiety in schizophrenia. *Schizophrenia Research: Cognition*, *4*, 1–3.

Addington, J., & Piskulic, D. (2013). Social cognition early in the course of the illness. In D. L. Roberts, & D. L. Penn (Eds.), *Social cognition in schizophrenia* (pp. 245–262). New York, NY: Oxford University Press.

Addington, J., Saeedi, H., & Addington, D. (2006a). Facial affect recognition: A mediator between cognitive and social functioning in psychosis? *Schizophrenia Research*, *85*(1), 142–150.

Addington, J., Saeedi, H., & Addington, D. (2006b). Influence of social perception and social knowledge on cognitive and social functioning in early psychosis. *The British Journal of Psychiatry*, *189*(4), 373–378.

Addington, J., Penn, D., Woods, S. W., Addington, D., & Perkins, D. O. (2008). Facial affect recognition in individuals at clinical high risk for psychosis. *The British Journal of Psychiatry*, *192*(1), 67–68.

Addington, J., Girard, T. A., Christensen, B. K., & Addington, D. (2010). Social cognition mediates illness-related and cognitive influences on social function in patients with schizophrenia-spectrum disorders. *Journal of Psychiatry and Neuroscience*, *35*(1), 49–55.

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, *60*, 693–716.

Allott, K. A., Rice, S., Bartholomeusz, C. F., Klier, C., Schlögelhofer, M., Schäfer, M. R., & Amminger, G. P. (2015). Emotion recognition in unaffected first-degree relatives of individuals with first-episode schizophrenia. *Schizophrenia Research*, *161*(2), 322–328.

Amminger, G. P., Schäfer, M. R., Klier, C. M., Schlögelhofer, M., Mossaheb, N., Thompson, A., ... Nelson, B. (2012a). Facial and vocal affect perception in people at ultra-high risk of psychosis, first-episode schizophrenia and healthy controls. *Early Intervention in Psychiatry*, *6*(4), 450–454.

Amminger, G. P., Schäfer, M. R., Papageorgiou, K., Klier, C. M., Schlögelhofer, M., Mossaheb, N., ... McGorry, P. D. (2012b). Emotion recognition in individuals at clinical high-risk for schizophrenia. *Schizophrenia Bulletin*, *38*(5), 1030–1039.

An, S. K., Kang, J. I., Park, J. Y., Kim, K. R., Lee, S. Y., & Lee, E. (2010). Attribution bias in ultra-high risk for psychosis and first-episode schizophrenia. *Schizophrenia Research*, *118*(1), 54–61.

Barkl, S. J., Lah, S., Harris, A. W., & Williams, L. M. (2014). Facial emotion identification in early-onset and first-episode psychosis: A systematic review with meta-analysis. *Schizophrenia Research*, *159*(1), 62–69.

Bartholomeusz, C. F., Allott, K., Killackey, E., Liu, P., Wood, S. J., & Thompson, A. (2013). Social cognition training as an intervention for improving functional outcome in first-episode psychosis: A feasibility study. *Early Intervention in Psychiatry*, *7*(4), 421–426.

Bediou, B., Asri, F., Brunelin, J., Krolak-Salmon, P., D'Amato, T. H., Saoud, M., & Tazi, I. (2007). Emotion recognition and genetic vulnerability to schizophrenia. *The British Journal of Psychiatry*, *191*(2), 126–130.

Bell, M., Bryson, G., & Lysaker, P. (1997). Positive and negative affect recognition in schizophrenia: A comparison with substance abuse and normal control subjects. *Psychiatry Research*, *73*(1), 73–82.

Bertrand, M. C., Sutton, H., Achim, A. M., Malla, A. K., & Lepage, M. (2007). Social cognitive impairments in first episode psychosis. *Schizophrenia Research*, *95*(1), 124–133.

Bliksted, V., Fagerlund, B., Weed, E., Frith, C., & Videbech, P. (2014). Social cognitive and neurocognitive deficits in first-episode schizophrenia. *Schizophrenia Research*, *153*(1), 9–17.

Bora, E., & Murray, R. M. (2014). Meta-analysis of cognitive deficits in ultra-high risk to psychosis and first-episode psychosis: Do the cognitive deficits progress over, or after, the onset of psychosis? *Schizophrenia Bulletin*, *40*(4), 744–755.

Bora, E., & Pantelis, C. (2013). Theory of mind impairments in first-episode psychosis, individuals at ultra-high risk for psychosis and in first-degree relatives of schizophrenia: Systematic review and meta-analysis. *Schizophrenia Research*, *144*(1), 31–36.

Bresnahan, M., Begg, M. D., Brown, A., Schaefer, C., Sohler, N., Insel, B., ... Susser, E. (2007). Race and risk of schizophrenia in a US birth cohort: Another example of health disparity? *International Journal of Epidemiology*, *36*(4), 751–758.

Brüne, M. (2005). "Theory of mind" in schizophrenia: A review of the literature. *Schizophrenia Bulletin*, *31*(1), 21–42.

Brüne, M., Özgürdal, S., Ansorge, N., von Reventlow, H. G., Peters, S., Nicolas, V., ... Lissek, S. (2011). An fMRI study of "theory of mind" in at-risk states of psychosis: Comparison with manifest schizophrenia and healthy controls. *NeuroImage*, *55*(1), 329–337.

Buck, B. E., Healey, K. M., Gagen, E. C., Roberts, D. L., & Penn, D. L. (2015). Social cognition in schizophrenia: Factor structure, clinical and functional correlates. *Journal of Mental Health*, 1–8.

Chan, R. C., Li, H., Cheung, E. F., & Gong, Q. Y. (2010). Impaired facial emotion perception in schizophrenia: a meta-analysis. *Psychiatry Research*, *178*(2), 381–390.

Comparelli, A., De Carolis, A., Corigliano, V., Romano, S., Kotzalidis, G. D., Campana, C., ... Girardi, P. (2011). Subjective disturbance of perception is related to facial affect recognition in schizophrenia. *The Journal of Nervous and Mental Disease*, *199*(10), 802–806.

Comparelli, A., Corigliano, V., De Carolis, A., Mancinelli, I., Trovini, G., Ottavi, G., ... Girardi, P. (2013). Emotion recognition impairment is present early and is stable throughout the course of schizophrenia. *Schizophrenia Research*, *143*(1), 65–69.

Corcoran, R., Cahill, C., & Frith, C. D. (1997). The appreciation of visual jokes in people with schizophrenia: A study of 'mentalizing' ability. *Schizophrenia Research*, *24*(3), 319–327.

Couture, S. M., Penn, D. L., & Roberts, D. L. (2006). The functional significance of social cognition in schizophrenia: A review. *Schizophrenia Bulletin*, *32*(Suppl. 1), S44–S63.

Couture, S. M., Penn, D. L., Addington, J., Woods, S. W., & Perkins, D. O. (2008). Assessment of social judgments and complex mental states in the early phases of psychosis. *Schizophrenia Research*, *100*(1–3), 237–241.

Daros, A. R., Ruocco, A. C., Reilly, J. L., Harris, M. S., & Sweeney, J. A. (2014). Facial emotion recognition in first-episode schizophrenia and bipolar disorder with psychosis. *Schizophrenia Research*, *153*(1), 32–37.

Dean, A. M., Goodby, E., Ooi, C., Nathan, P. J., Lennox, B. R., Scoriels, L., ... Barnes, A. (2013). Speed of facial affect intensity recognition as an endophenotype of first-episode psychosis and associated limbic-cortical grey matter systems. *Psychological Medicine*, *43*(3), 591–602.

Edwards, J., Pattison, P. E., Jackson, H. J., & Wales, R. J. (2001). Facial affect and affective prosody recognition in first-episode schizophrenia. *Schizophrenia Research*, *48*(2–3), 235–253.

Fett, A. K. J., Viechtbauer, W., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *35*(5), 573–588.

Fisher, M., Loewy, R., Hardy, K., Schlosser, D., & Vinogradov, S. (2013). Cognitive interventions targeting brain plasticity in the prodromal and early phases of schizophrenia. *Annual Review of Clinical Psychology*, *9*, 435–463.

Fornells-Ambrojo, M., & Garety, P. A. (2009). Understanding attributional biases, emotions and self-esteem in poor me paranoia: Findings from an early psychosis sample. *British Journal of Clinical Psychology*, *48*(Pt.2), 141–162.

Garety, P. A., & Freeman, D. (2013). The past and future of delusions research: From the inexplicable to the treatable. *The British Journal of Psychiatry*, *203*(5), 327–333.

Green, M. F., Penn, D. L., Bentall, R., Carpenter, W. T., Gaebel, W., Gur, R. C., ... Heinssen, R. (2008). Social cognition in schizophrenia: An NIMH workshop on definitions, assessment, and research opportunities. *Schizophrenia Bulletin*, *34*(6), 1211–1220.

Green, M. F., Bearden, C. E., Cannon, T. D., Fiske, A. P., Hellemann, G. S., Horan, W. P., ... Nuechterlein, K. H. (2012). Social cognition in schizophrenia, part 1: Performance across phase of illness. *Schizophrenia Bulletin*, *38*(4), 854–864.

Herbener, E. S., Hill, S. K., Marvin, R. W., & Sweeney, J. A. (2005). Effects of antipsychotic treatment on emotion perception deficits in first-episode schizophrenia. *American Journal of Psychiatry*, *162*(9), 1746–1748.

Herold, R., Feldmann, A., Simon, M., Tenyi, T., Kövér, F., Nagy, F., ... Fekete, S. (2009). Regional gray matter reduction and theory of mind deficit in the early phase of schizophrenia: A voxel-based morphometric study. *Acta Psychiatrica Scandinavica*, *119*(3), 199–208.

Hill, S. K., Sweeney, J. A., Hamer, R. M., Keefe, R. S., Perkins, D. O., Gu, H., ... Lieberman, J. A. (2008). Efficiency of the CATIE and BACS neuropsychological batteries in assessing cognitive effects of antipsychotic treatments in schizophrenia. *Journal of the International Neuropsychological Society*, *14*(2), 209–221.

Ho, K. K., Lui, S. S., Hung, K. S., Wang, Y., Li, Z., Cheung, E. F., & Chan, R. C. (2015). Theory of mind impairments in patients with first-episode schizophrenia and their unaffected siblings. *Schizophrenia Research*, *166*(1), 1–8.

Horan, W. P., Green, M. F., DeGroot, M., Fiske, A., Hellemann, G., Kee, K., ... Nuechterlein, K. H. (2012). Social cognition in schizophrenia, part 2: 12-month stability and prediction of functional outcome in first-episode patients. *Schizophrenia Bulletin*, *38*(4), 865–872.

Inoue, Y., Yamada, K., Hirano, M., Shinohara, M., Tamaoki, T., Iguchi, H., ... Kanba, S. (2006). Impairment of theory of mind in patients in remission following first episode of schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, *256*(5), 326–328.

Kay, S. R., Fiszbein, A., & Opfer, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276.

Kern, R. S., Penn, D. L., Lee, J., Horan, W. P., Reise, S. P., Ochsner, K. N., ... Green, M. F. (2013). Adapting social neuroscience measures for schizophrenia clinical trials, part 2: Trolling the depths of psychometric properties. *Schizophrenia Bulletin*, *39*(6), 1201–1210.

Kettle, J. W., O'Brien-Simpson, L., & Allen, N. B. (2008). Impaired theory of mind in first-episode schizophrenia: Comparison with community, university and depressed controls. *Schizophrenia Research*, *99*(1–3), 96–102.

Koelkebeck, K., Pedersen, A., Suslow, T., Kueppers, K. A., Arolt, V., & Ohrmann, P. (2010). Theory of mind in first-episode schizophrenia patients: Correlations with cognition and personality traits. *Schizophrenia Research*, *119*(1–3), 115–123.

Kohler, C. G., Turner, T. H., Bilker, W. B., Brensinger, C. M., Siegel, S. J., Kanes, S. J., ... Gur, R. C. (2003). Facial emotion recognition in schizophrenia: Intensity effects and error pattern. *American Journal of Psychiatry*, *160*(10), 1768–1774.

Kucharska-Pietura, K., David, A. S., Masiak, M., & Phillips, M. L. (2005). Perception of facial and vocal affect by people with schizophrenia in early and late stages of illness. *The British Journal of Psychiatry*, *187*, 523–528.

Kurtz, M. M., Gagen, E., Rocha, N. B., Machado, S., & Penn, D. L. (2016). Comprehensive treatments for social cognitive deficits in schizophrenia: A critical review and effect-size analysis of controlled studies. *Clinical Psychology Review*, *43*, 80–89.

Langdon, R., Connors, M. H., Still, M., Ward, P. B., & Catts, S. (2014a). Theory of mind and neurocognition in early psychosis: A quasi-experimental study. *BMC Psychiatry*, *14*, 316.

Langdon, R., Still, M., Connors, M. H., Ward, P. B., & Catts, S. V. (2014b). Theory of mind in early psychosis. *Early Intervention in Psychiatry*, *8*(3), 286–290.

Lee, S. Y., Bang, M., Kim, K. R., Lee, M. K., Park, J. Y., Song, Y. Y., ... An, S. K. (2015). Impaired facial emotion recognition in individuals at ultra-high risk for psychosis and with first-episode schizophrenia, and their associations with neurocognitive deficits and self-reported schizotypy. *Schizophrenia Research*, *165*(1), 60–65.

Leung, J. S. Y., Lee, T. M., & Lee, C. C. (2011). Facial emotion recognition in Chinese with schizophrenia at early and chronic stages of illness. *Psychiatry Research*, *190*(2–3), 172–176.

Mancuso, F., Horan, W. P., Kern, R. S., & Green, M. F. (2011). Social cognition in psychosis: Multidimensional structure, clinical correlates, and relationship with functional outcome. *Schizophrenia Research*, *125*(2–3), 143–151.

Mazza, M., Costagliola, C., Di Michele, V., Magliani, V., Pollice, P., Ricci, A., ... Galzio, R. J. (2007). Deficit of social cognition in subjects with surgically treated frontal lobe lesions and in subjects affected by schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, *257*(1), 12–22.

Mazza, M., Pollice, R., Pacitti, F., Pino, M. C., Mariano, M., Tripaldi, S., ... Roncone, R. (2012). New evidence in theory of mind deficits in subjects with chronic schizophrenia and first episode: Correlation with symptoms, neurocognition and social function. *Rivista di Psichiatria*, *47*(4), 327–336.

Mazza, M., Catalucci, A., Pino, M. C., Giusti, L., Nigri, A., Pollice, R., ... Gallucci, M. (2013). Dysfunctional neural networks associated with impaired social interactions in early psychosis: An ICA analysis. *Brain Imaging and Behavior*, *7*(3), 248–259.

McCleery, A., Horan, W. P., & Green, M. F. (2014). Social cognition during the early phase of schizophrenia. *Social cognition and metacognition in schizophrenia: Psychopathology and treatment approaches* (pp. 49–67). San Diego, CA: Academic Press.

Melle, I., Larsen, T. K., Haahr, U., Friis, S., Johannesen, J. O., Opjordsmoen, S., ... McGlashan, T. (2008). Prevention of negative symptom psychopathologies in first-episode schizophrenia: Two-year effects of reducing the duration of untreated psychosis. *Archives of General Psychiatry*, *65*(6), 634–640.

Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology*, *23*(3), 315–336.

Montreuil, T., Bodnar, M., Bertrand, M. C., Malla, A. K., Joober, R., & Lepage, M. (2010). Social cognitive markers of short-term clinical outcome in first-episode psychosis. *Clinical Schizophrenia & Related Psychoses*, *4*(2), 105–114.

Nuechterlein, K. H., Barch, D. M., Gold, J. M., Goldberg, T. E., Green, M. F., & Heaton, R. K. (2004). Identification of separable cognitive factors in schizophrenia. *Schizophrenia Research*, *72*(1), 29–39.

Olbert, C. M., Penn, D. L., Kern, R. S., Lee, J., Horan, W. P., Reise, S. P., ... Green, M. F. (2013). Adapting social neuroscience measures for schizophrenia clinical trials, part 3: Fathoming external validity. *Schizophrenia Bulletin*, *39*(6), 1211–1218.

Pantelis, C., Yücel, M., Bora, E., Fornito, A., Testa, R., Brewer, W. J., ... Wood, S. J. (2009). Neurobiological markers of illness onset in psychosis and schizophrenia: The search for a moving target. *Neuropsychology Review*, *19*(3), 385–398.

Pantelis, C., Wannan, C., Bartholomeusz, C. F., Allott, K., & McGorry, P. D. (2015). Cognitive intervention in early psychosis—Preserving abilities versus remediating deficits. *Current Opinion in Behavioral Sciences*, *4*, 63–72.
Penn, D. L., Corrigan, P. W., Bentall, R. P., Racenstein, J., & Newman, L. (1997). Social cognition in schizophrenia. *Psychological Bulletin*, *121*(1), 114–132.
Pinkham, A. E., Penn, D. L., Perkins, D. O., Graham, K. A., & Siegel, M. (2007). Emotion perception and social skill over the course of psychosis: A comparison of individuals "at-risk" for psychosis and individuals with early and chronic schizophrenia spectrum illness. *Cognitive Neuropsychiatry*, *12*(3), 198–212.
Pinkham, A. E., Penn, D. L., Green, M. F., Buck, B., Healey, K., & Harvey, P. D. (2014). The social cognition psychometric evaluation study: Results of the expert survey and RAND panel. *Schizophrenia Bulletin*, *40*(4), 813–823.
Piskulic, D., & Addington, J. (2011). Social cognition and negative symptoms in psychosis. *Psychiatry Research*, *188*(2), 283–285.
Reske, M., Habel, U., Kellermann, T., Backes, V., Shah, N. J., von Wilmsdorff, M., ... Schneider, F. (2009). Differential brain activation during facial emotion discrimination in first-episode schizophrenia. *Journal of Psychiatric Research*, *43*(6), 592–599.
Romero-Ferreiro, M. V., Aguado, L., Rodriguez-Torresano, J., Palomo, T., Rodriguez-Jimenez, R., & Pedreira-Massa, J. L. (2016). Facial affect recognition in early and late-stage schizophrenia patients. *Schizophrenia Research*, *172*(1), 177–183.
Savla, G. N., Vella, L., Armstrong, C. C., Penn, D. L., & Twamley, E. W. (2013). Deficits in domains of social cognition in schizophrenia: A meta-analysis of the empirical evidence. *Schizophrenia Bulletin*, *39*(5), 979–992.
Seiferth, N. Y., Pauly, K., Kellermann, T., Shah, N. J., Ott, G., Herpertz-Dahlmann, B., ... Habel, U. (2009). Neuronal correlates of facial emotion discrimination in early onset schizophrenia. *Neuropsychopharmacology*, *34*(2), 477–487.
Sergi, M. J., Rassovsky, Y., Widmark, C., Reist, C., Erhart, S., Braff, D. L., ... Green, M. F. (2007). Social cognition in schizophrenia: Relationships with neurocognition and negative symptoms. *Schizophrenia Research*, *90*(1–3), 316–324.
So, S. H. W., Tang, V., & Leung, P. W. L. (2015). Dimensions of delusions and attribution biases along the continuum of psychosis. *PloS One*, *10*(12), e0144558.
Sullivan, S., Lewis, G., Mohr, C., Herzig, D., Corcoran, R., Drake, R., & Evans, J. (2014). The longitudinal association between social functioning and theory of mind in first-episode psychosis. *Cognitive Neuropsychiatry*, *19*(1), 58–80.
Thompson, A., Papas, A., Bartholomeusz, C., Allott, K., Amminger, G. P., Nelson, B., ... Yung, A. (2012). Social cognition in clinical "at risk" for psychosis and first episode psychosis populations. *Schizophrenia Research*, *141*(2–3), 204–209.
Van Hooren, S., Versmissen, D., Janssen, I., Myin-Germeys, I., à Campo, J., Mengelers, R., ... Krabbendam, L. (2008). Social cognition and neurocognition as independent domains in psychosis. *Schizophrenia Research*, *103*(1–3), 257–265.
Vauth, R., Rüsch, N., Wirtz, M., & Corrigan, P. W. (2004). Does social cognition influence the relation between neurocognitive deficits and vocational functioning in schizophrenia? *Psychiatry Research*, *128*(2), 155–165.
Ventura, J., Wood, R. C., & Helleman, G. S. (2013). Symptom domains and neurocognitive functioning can help differentiate social cognitive processes in schizophrenia: A meta-analysis. *Schizophrenia Bulletin*, *39*(1), 102–111.
Ventura, J., Ered, A., Gretchen-Doorly, D., Subotnik, K. L., Horan, W. P., Hellemann, G. S., & Nuechterlein, K. H. (2015). Theory of mind in the early course of schizophrenia: Stability, symptom and neurocognitive correlates, and relationship with functioning. *Psychological Medicine*, *45*(10), 2031–2043.
Vohs, J. L., Lysaker, P. H., Francis, M. M., Hamm, J., Buck, K. D., Olesek, K., ... Breier, A. (2014). Metacognition, social cognition, and symptoms in patients with first episode and prolonged psychoses. *Schizophrenia Research*, *153*(1–3), 54–59.
Walter, H., Ciaramidaro, A., Adenzato, M., Vasic, N., Ardito, R. B., Erk, S., & Bara, B. G. (2009). Dysfunction of the social brain in schizophrenia is modulated by intention type: An fMRI study. *Social Cognitive and Affective Neuroscience*, *4*(2), 166–176.
Williams, L. M., Whitford, T. J., Nagy, M., Flynn, G., Harris, A. W., Silverstein, S. M., & Gordon, E. (2009). Emotion-elicited gamma synchrony in patients with first-episode schizophrenia: A neural correlate of outcomes. *Journal of Psychiatry & Neuroscience*, *34*(4), 303–313.
Yang, C., Zhang, T., Li, Z., Heeramun-Aubeeluck, A., Liu, N., Huang, N., ... Chen, F. (2015). The relationship between facial emotion recognition and executive functions in first-episode patients with schizophrenia and their siblings. *BMC Psychiatry*, *15*(1), 241.