Cognitive Neuropsychiatry, 2015
Vol. 20, No. 3, 198–221, http://dx.doi.org/10.1080/13546805.2014.999915

Routledge
Taylor & Francis Group

Observable Social Cognition – A Rating Scale: an interview-based assessment for schizophrenia

Kristin M. Healey[a], Dennis R. Combs[b], Clare M. Gibson[c], Richard S.E. Keefe[d], David L. Roberts[e] and David L. Penn[a,f]*

[a]*Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA;* [b]*Department of Psychology, University of Texas at Tyler, Tyler, TX, USA;* [c]*Research Service, Baltimore VA Healthcare System, Baltimore, MD, USA;* [d]*Department of Psychiatry, Duke University Medical Center, Durham, NC, USA;* [e]*Department of Psychiatry, University of Texas Health Science Center, San Antonio, TX, USA;* [f]*School of Psychology, Australian Catholic University, Melbourne, VIC, Australia*

(*Received 10 June 2014; accepted 13 December 2014*)

Introduction. Individuals with schizophrenia consistently show impairments in social cognition (SC). SC has become a potential treatment target due to its association with functional outcomes. An alternative method of assessment is to administer an observer-based scale incorporating an informant's "first hand" impressions in ratings.
Methods. The present study used the Observable Social Cognition: A Rating Scale (OSCARS) in 62 outpatients and 50 non-psychiatric controls (NPCs) to assess performance in domains of SC (e.g. emotion perception, theory of mind).
Results. The OSCARS demonstrated sufficient internal consistency and test-retest reliability. Construct validity was assessed through an exploratory factor analysis. Patient OSCARS indices were not significantly correlated with measures of SC with the exception of aggressive attributional style. Individuals with less impairment in SC reacted more aggressively to ambiguous situations. NPC OSCARS were significantly correlated with measures of theory of mind and attributional style. In a combined sample of patients and controls, six of eight items were significantly correlated with the SC task assessing the same domain, providing modest evidence of convergent validity. In patients, the OSCARS was significantly correlated with measures of functional outcome and neurocognition. Last, the OSCARS was found to be significantly associated with functional outcome after the influence of objective measures of SC was statistically removed.
Conclusions. The present study provides preliminary evidence that the OSCARS may be useful for clinicians in collecting data about patients' potential real-world SC deficits, in turn increasing the degree to which these impairments may be targeted in treatment.

Keywords: social cognition; schizophrenia; schizophrenia spectrum illness; measurement; functional outcome

Overview of social cognition (SC)

SC may be defined as a set of neurocognitive processes related to the understanding, recognition, processing and appropriate use of social stimuli in one's environment (Adolphs, 2009; Ochsner, 2008; Penn, Corrigan, Bentall, Racenstein & Newman, 1997). Individuals with schizophrenia consistently show impairments in SC across the following

*Corresponding author. Email: dpenn@email.unc.edu

primary domains: attributional style, theory of mind, emotion perception and associated underlying processes (Green, Olivier, Crawley, Penn, & Silverstein, 2005; Kohler, Walker, Martin, Healey, & Moberg, 2010; Penn, Sanna, & Roberts, 2008; Pijnenborg et al., 2009; Pinkham & Penn, 2006). SC has received considerable attention in the field of schizophrenia research over the past 10 years due to its relationship with poor functional outcomes (Brekke, Kay, Lee, & Green, 2005; Couture, Penn, & Roberts, 2006; Nuechterlein et al., 2004). Recent findings from a meta-analysis indicate that SC has a stronger relationship with functional outcome than neurocognition (Fett et al., 2011).

Problems with measuring SC

Given the importance of SC to social functioning, it is critical to utilise valid and reliable measures to enhance our understanding of these constructs. Current measures often have important methodological issues that limit the utility of SC as a viable treatment target. First, SC tasks' psychometric properties are often not well established (Bora, Yucel, & Pantelis, 2009; Pinkham et al., 2014; Yager & Ehmann, 2006). And second, some of the current SC tasks have significant conceptual and measurement-related overlap (Green et al., 2008). For example, the Eyes task prompts subjects to label pictures of eyes with a word that best categorises their interpretation of the person's experience (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997). This task is meant to assess theory of mind, but likely involves aspects of emotion perception. Such problems call for the supplementation of existing measures with novel methods of assessing SC.

Observer-based scales

An alternative approach to measuring SC deficits is to administer an observer-based rating scale. This method was used for the Schizophrenia Cognition Rating Scale (SCoRS), an interview-based measure that considers informant reports, or information from individuals that had the most regular contact with the patient in everyday situations. The SCoRS was found to be a valid assessment of cognition, as global ratings were significantly correlated with composite scores of cognitive performance, but only when informant data were included (Keefe, Poe, Walker, Kang, & Harvey, 2006). Ventura, Cienfuegos, Boxer, and Bilder (2008; Ventura et al., 2010) developed similar scales incorporating informant reports, the Clinical Global Impression of Cognition in Schizophrenia (CGI-CogS) and subsequently the Cognitive Assessment Interview from a subset of SCoRS and CGI-CogS items. Both were found to be valid assessments of cognition (Ventura et al., 2010).

The present study: aims and hypotheses

The present study evaluated the psychometric characteristics of a new observer-based rating scale of SC incorporating informant ratings, the Observable Social Cognition: A Rating Scale (OSCARS). First, this study evaluated the internal consistency and the test-retest reliability of the OSCARS over an approximate one-week period. Second, the construct validity was investigated through an exploratory factor analysis (EFA) of the OSCARS. Construct validity was also assessed with group comparisons and analyses of diagnostic sensitivity. Third, the convergent validity of the scale was examined via the relationship between the OSCARS and measures of emotion perception, theory of mind, attributional style and jumping to conclusions. Fourth, external validity was explored through investigating the relationship between the OSCARS and measures of social skill and

social/role functioning. Fifth, it is expected that IQ and cognition will be moderately associated with ratings on the OSCARS, which will provide evidence of discriminant validity. And sixth, predictive validity was explored through investigating whether the OSCARS will explain more variance in functional outcome than laboratory-based measures of SC.

Methods

Participants

Sixty-two individuals aged 25–60 with schizophrenia spectrum disorders (schizoaffective = 35; schizophrenia = 27) and without current substance use problems were recruited through a university-based outpatient clinic and mental health centres in the Chapel Hill area. Individuals were participating in a study of social cognition and interaction training (SCIT), a 20–24 week, manual-based group intervention that targets dysfunctional SC processes (Roberts et al., 2014). Laboratory-based measures of SC were selected based on domains of SC that are targeted in SCIT, and thus which would be most likely to show a treatment effect. Within each domain, measures were selected that are commonly used in patients with schizophrenia spectrum illnesses. Screening procedures involved administration of the Wechsler Abbreviated Scales for Intelligence (WASI; Whitmyre & Pishkin, 1958) to rule out any individuals with possible mental retardation (IQ < 70). Individuals diagnosed with a major nervous system disorder (e.g., seizure disorder) were also excluded from participation. Participants were required to endorse a mild (2) or greater level of social impairment as determined by a subgroup of interaction items from the Social Functioning Scale (Birchwood, Smith, Cochrane, Wetton, & Copestake, 1990).

Diagnoses were assessed through review of participants' medical charts and confirmed with items from the psychotic disorders section of the Structured Clinical Interview for DSM-IV – Patient Edition (Werner, 2001).

Fifty non-psychiatric controls (NPCs) aged 18–65 were recruited through flyers and craigslist.org postings. Controls must have reported no history of mental illness and no first-degree relatives with a psychotic disorder, bipolar disorder or autism.

Development of OSCARS

OSCARS item generation

The OSCARS is an 8-item, interview-based assessment of SC in outpatients with schizophrenia (see Appendix). These items were developed by the study's principal investigators (Drs Penn, Combs, and Roberts) to broadly assess the SC domains of theory of mind, emotion perception, cognitive rigidity, jumping to conclusions and attributional style. These areas were selected because they have shown consistent deficits in patients with schizophrenia. The initial pool included 11 items that were reviewed for validity by five experts in the field of SC: Drs Patrick Corrigan (Illinois Institute of Technology), Allen Fenigstein (Kenyon College), Daniel Freeman (Oxford University), William Horan (UCLA) and Kim Mueser (Boston University). Experts rated each item on a 1–5 scale (1 = lowest level of validity and 5 = greatest level of validity). Items that reached an average rating of 3 or above were retained. Three items were removed because they were not considered to be valid indicators of SC, but rather of social skill, self-awareness and insight.

Each OSCARS item is comprised of a question probing a SC construct followed by general example behaviours that reflect impairment in that domain. Each item is scored

on a 7-point Likert-type scale, higher ratings indicating greater observed impairment. Anchor points were created for four levels (1, 3, 5, 7), and captured degree of impairment (severity, frequency).

OSCARS administration

The OSCARS can be administered one of two ways, either as a semi-structured interview with the subject or as an informant-based questionnaire. For all participants in the present study, the OSCARS was administered as a semi-structured interview (n = 112). Both methods take approximately 15–20 minutes to administer and rate. The informant was provided with a copy of the instrument and directly selected each rating on the 7-point scale, utilising the anchors provided. Thus, informant ratings were based solely on that individual's report, specifically regarding their interaction with and knowledge of the individual. For a subset of subjects (n = 39), complete administration of the OSCARS generated an additional interviewer rating. The interviewer rating is an integrated rating that considered the information provided by the informant and permitted the interviewer to agree or disagree with the informant's rating. Informant and interviewer ratings were significantly correlated (r = .94, p < .001), thus all subsequent analyses use informant ratings.

We aimed to interview the informant who had the most regular contact with the subject in everyday situations. In this study, informants held a variety of roles: first-degree family members (n = 29), friends (n = 8), significant others (n = 6), roommates (n = 4), other family members (n = 3), social workers (n = 3), clubhouse staff (n = 3), therapists (n = 2), supervisors (n = 2), pastor (n = 1) and job counsellor (n = 1; n = 62 total). Healthy control informants had the following roles: first-degree family members (n = 10), friends (n = 13), significant others (n = 23), roommates (n = 3) and other family members (n = 1; n = 50 total).

Social cognitive measures

Emotion perception

The Face Emotion Discrimination Task (FEDT; range 0–30) and the Face Emotion Identification Task (FEIT; range 0–19) were used to measure emotion perception (Kerr & Neale, 1993). On these two measures, performance is indexed as the total number of correct items.

The protocol was later supplemented with the Penn Emotion Recognition Test (ER40; Kohler et al., 2003; range 0–40). Performance is indexed as the total number of correct items (Table 1).

Theory of mind

Both the Hinting Task (Corcoran, 2003; range 0–20) and The Awareness of Social Inference Test, Part 2 (social inference-minimal; TASIT; McDonald, Flanagan, Rollins, & Kinch, 2003; range 0–60) were used to measure theory of mind. The total number of items correct indexes performance.

Attributional style

The Ambiguous Intentions Hostility Questionnaire (AIHQ; Combs, Penn, Wicher, & Waldheter, 2007) was used to measure attributional style. Higher ratings indicate greater

Table 1. Psychometric characteristics of measures of SC.

	Reliability	Convergent validity	Criterion validity
Emotion perception			
FEDT	SCZ α = .32; HC data unavailable	FEDT, FEIT, and ER40 were significantly intercorrelated with *r*'s ranging from .36–.67 ($p < .05$).	Emotion perception tasks not significantly associated with any measures of functioning.
FEIT	SCZ α = .38; HC α = .84		
ER40	Unavailable in present sample		
Theory of mind			
TASIT	SCZ α = .55; HC data unavailable	Hinting task and TASIT were significantly correlated ($r = .45$, $p < .001$).	Theory of mind tasks not significantly associated with any indices of functioning.
Hinting task	SCZ α = .46; HC α = .70		
Attributional style			
AIHQ Blame Index	SCZ α = .86; HC α = .84	In HC, Blame index was associated with the ER40 ($r = -.30$, $p < .05$) and TASIT total ($r = -.42$, $p < .01$). In SCZ, FEDT was associated with AIHQ Blame ($r = -.25$, $p < .05$) and AIHQ Aggression ($r = 0.27$, $p < .05$).	AIHQ Blame subscale was correlated with the QLS: Social index ($r = -.27$, $p < .05$). Not associated with other measures of functioning.
AIHQ Hostility Bias	SCZ α = .43; HC α = .43		
AIHQ Aggression Index	SCZ α = .55; HC α = .26		
Jumping to conclusions			
Beads task	N/A	Not significantly correlated with any other measures of SC.	Beads was significantly associated with QLS: Social index ($r = .26$, $p < .05$).

aggression (range 5–25), hostility (range 5–25) and blame (range 15–80) biases. Coders were trained to inter-rater reliability of intraclass correlation coefficient (ICC) > .70 against a gold-standard rater criterion.

Probabilistic reasoning (jumping to conclusions)

The "beads in the jar" task (Dudley, John, Young, & Over, 1997a, 1997b; range 0–30) was used to assess jumping to conclusions. Greater number of beads selected before a decision is made indicates lower likelihood of jumping to conclusions. The range of beads selected for both patients and healthy controls was 1–20.

Functional measures

Social skill and social functioning

The Social Skills Performance Assessment (SSPA; Patterson, Moscona, McKibbin, Davidson, & Jeste, 2001; range 9–45 per role-play) was used to assess social skill/social functioning. Lower ratings indicate greater social skill impairment. Scores from the two SSPA role-plays were correlated ($r = .56$, $p < .001$), and so were collapsed to create a

total SSPA score (range 18–90). Intraclass correlations were computed and all were greater than 0.80 for all subscales.

The Global Social Functioning Scale (GSFS; Cornblatt et al., 2007; range 1–10) was used to measure social functioning. The GSFS yields a single global social/interpersonal functioning score between 0 and 10, with lower scores indicating greater impairment. Trained research clinicians determined the score based on information from informant report.

The Role Functioning Scale (RFS; McPheeters, 1984; range 4–28), a 4-item semi-structured interview, measures four major domains of everyday functioning. The RFS was conducted as an informant-based interview. Each item is rated on a scale of 1–7, higher ratings indicating greater functioning.

The Quality of Life Scale – Social (QLS-S; range 0–48) and Work (QLS-W; range 0–24; Heinrichs, Hanlon, & Carpenter, 1984) comprises eight- and four-item subscales, respectively. The QLS is an interview-based measure. Trained research clinicians determined the score based on information from the participant's responses

Intelligence quotient

The WASI was used to measure IQ, which consisted of administration of Matrix Reasoning and Vocabulary subtests.

Cognition

The SCoRS is an interview-based measure of cognition (Keefe et al., 2006). The interviewer global rating was utilised because it has the highest correlation with indices of functioning (Keefe et al., 2006). Each global rating is coded on a scale of 1–10, higher ratings indicating greater cognitive impairment.

Symptoms

The Positive and Negative Syndrome Scale (PANSS: Kay, Fiszbein, & Opler, 1987) was used to assess symptomatology. Higher scores indicate more severe symptoms.

Procedure

Study protocol was administered under the supervision of the principal investigator David L. Penn. All research assistants completed comprehensive training on administration of study measures prior to working with participants. Raters were required to achieve acceptable levels of inter-rater reliability (ICCs and Kappas > .80) on all interview-based measures. Raters were not blinded to group.

The OSCARS was administered at baseline and then again 7–10 days later to evaluate test-retest reliability (mean = 9.36, SD = 3.04). The same informant was interviewed at both baseline and retest for all subjects with complete retest data (n = 47). It should be noted that by retest, patients in the treatment group had begun weekly SCIT training. However, retesting occurred during introductory sessions (1–2), which are associated with minimal expected improvements in SC.

Results

Data analyses were performed using SPSS version 20 and Comprehensive Exploratory Factor Analysis version 3.04. Statistical significance was defined as $p < .05$.

Sample characteristics and descriptive statistics

There were no statistically significant differences between patient psychiatric control and NPC groups in baseline demographic variables with the exception of participant education and IQ (Table 2), which were later included as covariates. Table 3 displays descriptive statistics for OSCARS, SC measures and functional outcome measures.

Reliability analyses

The internal consistency of the OSCARS (Cronbach's alpha) was .80 in patients and .78 in controls. Test-retest reliability of the eight OSCARS items ranged from .50 to .70 (mean = .62, SD = .07). OSCARS total score test-retest reliability was .86 (n = 47; patients only).

Validity analyses

The construct validity of the OSCARS was evaluated via a factor analysis in patients (n = 62) and controls (n = 50) separately. An EFA examined whether the OSCARS loads on separable factors. The factor structure was determined by a preliminary examination of a scree plot and further investigated with a chi-square test and model fit indices. Maximum likelihood extraction method was used because it generally provides better estimates than other approaches (Fabrigar, Wegener, MacCallum, & Strahan, 1999).

Table 2. Demographic and clinical characteristics.

	Schizophrenia (N = 62)		Controls (N = 50)		Test statistics	
	n	M (SD)	n	M (SD)	t, X^2 (df)	p-value
Age	62	39.58 (11.47)	50	39.86 (9.85)	t = .14 (110)	.89
Education						
Participant	62	12.26 (1.21)	50	13.40 (1.18)	t = 5.02 (110)	.00
Mother	56	12.66 (2.37)	48	12.67 (1.92)	t = −.01 (102)	.99
Father	47	12.96 (2.65)	31	12.87 (1.59)	t = −.16 (76)	.87
WASI (IQ)	62	99.74 (15.28)	50	110.80 (15.00)	t = 3.84 (110)	.00
Age of first hospitalisation	62	22.71 (7.89)				
Number of hospitalisations	61	6.31 (6.65)				
PANSS symptoms						
Positive	62	16.32 (5.16)				
Negative	62	14.77 (4.44)				
General	62	33.65 (8.64)				
Total	62	65.66 (13.10)				
Sex (% male)	66.13		66.00		X^2 = .00 (df = 1)	1.00
Race/ethnicity						
Caucasian (%)	64.52		68.00			
African-American (%)	35.48		32.00		X^2 = .15 (df = 1)	.84
Hispanic/Latino						
Hispanic (%)	5.00		2.00		X^2 = .70 (df = 1)	.62

Table 3. Descriptive statistics for OSCARS, social cognitive measures and functional outcome measures.

	Schizophrenia			Controls			Test statistics	
	n	Mean (%)	SD (%)	*n*	Mean (%)	SD (%)	*t* (df)	*p*-value
OSCARS total	62	24.10	8.31	50	13.86	5.68	−7.42 (110)	.00
Emotion perception								
FEDT	62	25.00 (83.00)	2.22 (7.40)	50	25.28 (84.27)	2.18 (7.27)	.67 (110)	.50
FEIT	62	12.23 (64.37)	2.66 (14.00)	50	13.94 (73.37)	2.24 (7.27)	3.63 (110)	.00
ER40	28	30.43 (76.08)	5.47 (13.68)	49	33.16 (82.90)	2.71 (6.78)	2.93 (75)	.00
Theory of mind								
TASIT	62	47.35 (78.92)	7.16 (11.93)	50	53.78 (89.63)	5.12 (8.53)	5.34 (110)	.00
Hinting task	62	14.81 (74.05)	3.01 (15.05)	50	16.90 (84.5)	2.61 (13.05)	3.89 (110)	.00
Attributional style								
AIHQ Blame Index	62	41.45	13.39	50	36.60	10.84	−2.07 (110)	.04
AIHQ Hostility Bias	62	10.71	3.05	50	8.44	2.48	−4.25 (110)	.00
AIHQ Aggression Index	62	8.95	1.94	50	10.28	1.83	3.70 (110)	.00
Jumping to conclusions								
Beads task	62	8.06	5.29	50	8.60	4.88	.55 (110)	.58
Functioning								
GSFS	61	5.98	1.15	–	–	–	–	–
SSPA1: Total	60	28.40	4.63					
SSPA2: Total	60	28.38	5.41					
RFS: Working productivity	61	4.61	1.61	–	–	–	–	–
RFS: Independent living	61	5.38	1.34	–	–	–	–	–
RFS: Immediate social	61	5.51	.98	–	–	–	–	–
RFS: Extended social	61	5.05	1.41	–	–	–	–	–

Table 3 (*Continued*)

Schizophrenia				Controls			Test statistics	
	n	Mean (%)	SD (%)	*n*	Mean (%)	SD (%)	*t* (df)	*p*-value
Role of functionality total	61	20.54	3.70	–	–	–	–	–
QLS: Social	62	25.04	8.96	–	–	–	–	–
QLS: Work	62	14.32	4.78	–	–	–	–	–
QLS: Total	62	39.36	11.59					
Neurocognition								
SCoRS	62	4.92	2.48	–	–	–	–	–

FEDT, Face Emotion Discrimination Task; FEIT, Face Emotion Identification Task; ER40, Emotion Recognition; TASIT, The Awareness of Social Inference Test; AIHQ, Ambiguous Intentions Hostility Questionnaire; GSFS, Global Social Functioning Scale; SSPA, Social Skills Performance Assessment (1/2 denote role-play number); RFS, Role Functioning Scale; QLS, Quality of Life Scale; SCoRS, Social Cognition Rating Scale.

Crawford-Ferguson Quartimax, oblique rotation was selected because the factors are likely intercorrelated.

In participants with schizophrenia, a two-factor solution was the model of best fit. The root mean square error of approximation (RMSEA) was within the range of reasonable fit at .07 [confidence interval (CI): .00–.15; Browne, Cudeck, Bollen, & Long, 1993]. The Tucker-Lewis Index (TLI) was also adequate at .93 (Hu & Bentler, 1999). There was consensus between the scree plot and model fit for the selection of a two-factor model. Items were assigned to one factor depending on the magnitude of each factor loading (Table 4). The first factor contained high loadings for questions probing attributional style (2), jumping to conclusions (3) and cognitive rigidity (4, 5). Factor 1 was labelled "Social Cognitive Bias", as it appears to assess SC behavioural indicators of impulsivity, hostility and rigidity. The second factor contained high loadings for questions probing theory of mind (6, 7, 8) and emotion perception (1). Factor 2 was labelled "Social Cognitive Ability", as items share content involving perceptual and reasoning abilities. Item 6 (theory of mind) did not clearly load on one factor, thus it was retained on factor 2 with other items assessing theory of mind.

Factor scores were computed by summing OSCARS raw item scores that correspond to each factor. The factors were moderately intercorrelated with one another (r = .36, p < .05). Test-retest reliability was .87 for factor 1 and .85 for factor 2.

In controls, a three-factor solution was determined to be the model of best fit. The RMSEA was within the range of close fit at .04 (CI: .00–.19; Browne et al., 1993). The TLI indicates excellent model fit at .98 (Hu & Bentler, 1999). There was consensus between the scree plot and model fit for the selection of a three-factor model. Factors 1 and 2 in healthy controls are very similar to factors 1 and 2 in individuals with schizophrenia, and thus factor naming is consistent. The third factor contains high loadings for questions probing cognitive rigidity (4) and theory of mind (6). Factor 3 was labelled "Social Cognitive flexibility" as items assess flexibility in social situations and subtle theory of mind ability. The factors were moderately intercorrelated with one another (r = .36, p < .01).

Regarding construct validity, individuals with schizophrenia had significantly greater deficits on the OSCARS than NPCs, $F(1,108)$, p < .001; Table 3, after controlling for IQ

Table 4. Factor analysis of the OSCARS.

	Factor 1: social cognitive bias	Factor 2: social cognitive ability		Factor 1: social cognitive bias	Factor 2: social cognitive ability	Factor 3: social cognitive flexibility
	Schizophrenia			Controls		
Q1-EP	.33	**.43**	Q1-EP	−.01	**.75**	.03
Q2-AS	**.75**	−.16	Q2-AS	**.37**	−.26	.30
Q3-JTC	**.75**	.01	Q3-JTC	**.98**	−.03	.05
Q4-CR	**.82**	−.01	Q4-CR	.05	.00	**.98**
Q5-CR	**.50**	.36	Q5-CR	**.55**	.34	−.17
Q6-ToM	.28	**.21**	Q6-ToM	−.08	.35	**.51**
Q7-ToM/Emp	.32	**.47**	Q7-ToM/Emp	.10	**.43**	.34
Q8-ToM	−.08	**.97**	Q8-ToM	.21	**.37**	.18

Note: Bolded values indicate item loads on the factor titled above.
EP, emotion perception; AS, attributional style; JTC, jumping to conclusions; CR, cognitive rigidity; ToM, theory of mind; Emp, empathy.

and education. To assess diagnostic sensitivity, we conducted receiver operating characteristic (ROC) analyses to evaluate the potential for the OSCARS to be used as a diagnostic tool. A value of 1.0 indicates perfect diagnostic prediction and .50 indicates a level of chance. ROC analyses on OSCARS total scores indicated a high area under the curve estimate of .85 (95% CI = .78–.92; $p < .001$) in differentiating between individuals with schizophrenia and healthy controls. The optimal cut-off point suggested by the Youden Index was an OSCARS total score of 17 (sensitivity = .71, specificity = .78). Thus, anyone scoring higher than this cut-off may be considered scoring in the schizophrenia spectrum range.

In regard to convergent validity (Table 5), in the schizophrenia sample, OSCARS total and SC Ability (factor 2) scores were significantly negatively associated with AIHQ aggression index scores, indicating that higher aggressive attribution ratings (i.e. increased tendency to report acting aggressively in ambiguous situations) are correlated with less impairment in SC.

In controls, OSCARS total was significantly correlated with TASIT total score, indicating greater theory of mind performance is associated with less impairment in real-world SC. SC Bias (factor 1) was significantly correlated with AIHQ hostility bias; greater real-world SC impairment was associated with greater hostile attribution biases in ambiguous situations. Correlations between (1) hinting task and SC Bias (factor 1) and (2) beads task and SC Flexibility (factor 3) approached statistical significance.

To further assess convergent validity, a series of correlational analyses were conducted at the item level in the combined sample ($n = 112$). We sought to investigate the relationship between specific OSCARS items (e.g., Item 2 assessing Attributional style) and SC test scores meant to assess the same or closely related domains (e.g., AIHQ indices, Attributional style). Item 1 (emotion perception) was not significantly correlated with SC measures of emotion perception (FEDT: $r = .06$, $p = .56$; FEIT: $r = .05$, $p = .64$; ER40: $r = -.18$, $p = .12$). Item 2 (attributional style) was significantly correlated with the AIHQ Blame Index ($r = .20$, $p < .05$) and AIHQ Hostility Bias ($r = .20$, $p < .05$), but not the AIHQ Aggression Index ($r = -.16$, $p = .09$). Item 3 (jumping to conclusions) was not significantly correlated with the beads task ($r = -.15$, $p = .11$). Cognitive rigidity is thought to underlie domains of theory of mind and attributional style (Penn et al., 2008), thus we correlated items 4 and 5 with measures of theory of mind (Hinting Task and TASIT) and attributional style (AIHQ). Items 4 and 5 were significantly correlated with both theory of mind measures, with a range of correlations between −.23 and −.41 ($p < .01$). Item 4 was significantly correlated in the expected direction with AIHQ Blame ($r = .21$; $p < .05$) and AIHQ Hostility ($r = .28$, $p < .01$). However, consistent with the convergent validity findings in the schizophrenia group, AIHQ Aggression was significantly associated with item 4 ($r = -.23$, $p < .05$) and item 5 ($r = -.30$, $p < .01$) such that higher aggressive attribution ratings are correlated with less impairment in SC. Item 5 was associated with AIHQ Hostility ($r = .29$; $p < .01$), but not AIHQ Blame. Last, items 6, 7 and 8 (theory of mind/empathy) were significantly correlated with the TASIT ($r = -.21$ to $-.24$, $p < .05$), but not with the hinting task ($r = .02$ to $-.14$, $p > .13$).

In regard to external validity in the schizophrenia sample (Table 6), OSCARS total and SC Bias (factor 1) were significantly associated with GSFS scores such that less impairment in SC was associated with higher global social functioning. Several OSCARS indices were significantly correlated with RFS Working productivity and Independent Living subscales; greater productivity and independence were associated with less

Table 5. Convergent validity: correlations between OSCARS total informant score and OSCARS factor scores with measures of SC (n = 62).

	OSCARS total	Factor 1: social cognitive bias	Factor 2: social cognitive ability	OSCARS total	Factor 1: social cognitive bias	Factor 2: social cognitive ability	Factor 3: social cognitive flexibility
		Schizophrenia			Controls		
Emotion perception							
FEDT	−.02	−.11	.09	−.17	−.21	−.12	−.08
FEIT	−.02	−.04	.01	−.03	−.08	−.06	.13
ER40	.00	.16	−.20	.03	.08	.12	−.20
Theory of mind							
TASIT	−.14	−.13	−.10	**−.40****	**−.39****	−.24[#]	**−.37****
Hinting task	.07	−.03	.15	−.22	−.28[#]	−.15	−.07
Attributional style							
AIHQ Blame Index	.08	.10	.03	.00	.13	−.15	.07
AIHQ Hostility Bias	.11	.03	.16	.12	**.34***	−.15	.20
AIHQ Aggression Index	−.24[#]	−.14	**−.27***	−.12	.05	−.21	−.08
Jumping to conclusions							
Beads task	−.12	−.08	−.08	−.23	−.15	−.17	−.27[#]

Note: Bolded items indicate clinical significance.
FEDT, Face Emotion Discrimination Task; FEIT, Face Emotion Identification Task; ER40, Emotion Recognition; TASIT, The Awareness of Social Inference Test; AIHQ, Ambiguous Intentions Hostility Questionnaire.
*$p < 0.05$; **$p < 0.01$; [#]$p < .09$.

Table 6. Schizophrenia participants' external validity: correlations between OSCARS indices and measures of functional outcome.

	OSCARS total	Factor 1: social cognitive bias	Factor 2: social cognitive ability
GSFS	**−.27***	**−.30***	−.13
SSPA1: Total	−.07	−.09	−.03
SSPA2: Total	−.25#	−.23#	−.18
SSPA Total	−.19	−.19	−.12
RFS: Working productivity	**−.39****	**−.34****	**−.29***
RFS: Independent living	**−.28***	**−.27***	−.20
RFS: Immediate social	−.13	−.16	−.06
RFS: Extended social	−.21	−.24#	−.09
Role of functionality total	**−.38****	**−.38****	−.25#
QLS: Social	−.02	−.11	.10
QLS: Work	−.24#	−.20	−.20
QLS: Total	−.11	−.17	−.01

Note: Bolded items indicate clinical significance.
GSFS, Global Social Functioning Scale; SSPA, Social Skills Performance Assessment (1/2 denote role-play number); RFS, Role Functioning Scale; QLS, Quality of Life Scale.
*$p < 0.05$; **$p < 0.01$; #$p < .08$.

impairment in SC. OSCARS total and SC Bias (factor 1) scores were significantly associated with role functioning total scores; greater functionality was associated with lower deficits in SC. Additionally, several correlations approached statistical significance and were in the expected direction (see Table 6).

Discriminant validity was explored through computing correlations between OSCARS total or factor scores and interview-based measures of neurocognition in the schizophrenia sample only. All OSCARS indices were significantly associated with the SCoRS, including OSCARS total ($r = .67$, $p < .000$), SC Bias (factor 1; $r = .54$, $p < .000$), and SC Ability (factor 2; $r = .57$, $p < .000$). WASI Full Scale IQ was not significantly correlated with the OSCARS. The correlation between the WASI and SC Bias (factor 1) approached significance ($r = -.23$, $p = .069$). OSCARS was correlated with an observational index of cognition, but not a standardised IQ test score. In addition, there were no significant correlations between OSCARS total or factor scores and PANSS subscales. The range of correlations was −.10 to .20.

Predictive validity in the schizophrenia sample was explored through hierarchical regression to determine if the OSCARS total score explains more variance in functional outcome than laboratory-based measures of SC. Measures of functioning that were found to be significantly associated with OSCARS indices were entered as the dependent variable, each conducted as a separate analysis. For each analysis, predictor variables were entered into the model in the following order: (1) all indices of laboratory-based SC and (2) OSCARS total score. All indices of SC were entered as raw scores with the exception of the emotion perception indices (ER40, FEDT, FEIT), which were standardised and combined to create a composite index due to incomplete ER40 data collection ($n = 28$). Analyses included all individuals with available functional data ($n = 61$). The OSCARS was found to be significantly associated with real-world functioning after the effect of laboratory-based SC measures was statistically removed, as indicated by change in R^2, for the following indices: GSFS, $F = 4.59$, df = 1,50, $p <$

Table 7. Prediction of indices of functioning: regression based on objective measures of SC and OSCARS total score.

	R^2	F	df	P
Hierarchical regression predicting GSFS				
Objective measures of SC	.105	.667	9,51	.735
OSCARS Total Score	.181	4.591	1,50	**.037***
Hierarchical regression predicting RFS working productivity				
Objective measures of SC	.236	1.749	9,51	.102
OSCARS Total Score	.397	13.398	1,50	**.001****
Hierarchical regression predicting RFS independent living				
Objective measures of SC	.230	1.695	9,51	.114
OSCARS Total Score	.276	3.128	1,50	.083^
Hierarchical regression predicting RFS total				
Objective measures of SC	.142	.939	9,51	.501
OSCARS Total Score	.298	11.066	1,50	**.002****

Note: Bolded items indicate clinical significance.
*$p < 0.05$; **$p < 0.01$; ^$p < .10$.

.05; RFS Working Productivity, $F = 13.40$, df = 1,50, $p < .01$; and RFS Total, $F = 11.07$, df = 1,50, $p < .01$; Table 7. Further, the OSCARS showed trend level significance with RFS Independent living after accounting for variance from laboratory-based measures of SC, $F = 3.13$, df = 1,50, $p = .08$.

Exploratory analyses

Individuals with schizophrenia identified a nearly equal number of first-degree family members ($n = 29$) and other individuals ($n = 33$) as informants. Exploratory analyses were conducted to investigate potential differences in OSCARS ratings as a function of informant status. First-degree family members rated participants as having significantly greater SC deficits on the OSCARS than other informants, $F(1,60)$, $p = .008$; first-degree family mean = 26.72, SD = 7.89; other individuals mean = 21.79, SD = 5.68. Similar analyses were not conducted in the control group because of uneven sample size (first-degree family members, $n = 10$; other individuals, $n = 40$). However, healthy controls identified a nearly equal number of significant others ($n = 22$) and other individuals ($n = 28$), thus potential differences in OSCARS ratings as a function of informant were explored. There were no significant differences between groups, $F(1,48)$, $p = .732$; significant other mean = 13.55, SD = 4.48; other individuals mean = 14.11, SD = 6.54.

Discussion

The results of the present study indicate that the OSCARS is a psychometrically reliable, easily administered, observer-based measure of SC. The OSCARS had adequate test-retest reliability and internal consistency. Exploratory factor analyses yielded interpretable factors in both patient and healthy control data. The OSCARS displayed evidence of construct validity, as OSCARS total scores: (1) were significantly different between groups in the expected direction and (2) adequately differentiated between patients and controls in ROC analyses. OSCARS indices displayed weak evidence of convergent

validity with measures of SC. Correlational analyses of individual OSCARS items with measures in respective SC domains provided mild evidence of convergent validity. Finally, OSCARS indices were significantly correlated with various functional outcome measures.

The OSCARS total and factor scores did not show impressive convergent validity in patients with schizophrenia; they were not significantly associated with any measures of SC in the expected direction. Specifically, SC Ability (factor 2) was significantly negatively correlated with the AIHQ Aggression Index, indicating individuals with less SC impairment report more aggressive responses to hypothetical ambiguous situations. Correlations between the OSCARS and AIHQ Hostility Bias were non-significant, suggesting that aggressive reactions were not preceded by hostile biases. This is contrary to foundational work on attributional biases in aggressive boys, which posits aggressive behaviours occur as a result of systematic hostile biases (Dodge, 2006). However, individuals with serious mental illness are often targets of stigma, thus participants with higher SC may expect social situations to be more stigmatising, and respond to them in a more reactive/automatic manner.

The absence of significant associations may be due to error variance in the validity of informant report. Sabbag et al. (2011) found that high contact clinicians provided ratings of patients' real-world functioning that were more closely related to objective indices than the ratings of friends or family members. The present study used a heterogeneous group of informants, which may have obfuscated potentially significant correlations.

The lack of evidence of convergent validity of the OSCARS may also be related to the difficulty of capturing true score variance when conducting separate group analyses. Thus, the groups were collapsed to conduct item level correlational analyses. In a combined sample of patients and controls, six of eight OSCARS items were significantly correlated with the SC task assessing the same domain, providing modest evidence of convergent validity. Items assessing attributional style, theory of mind and cognitive rigidity were significantly correlated with respective SC tasks. With the exception of the AIHQ Aggression Index, all correlations were in the expected direction, meaning poor performance on SC tests was correlated with greater observed SC deficits on the OSCARS. This suggests that individual OSCARS items may be appropriately tapping into the posed SC domain as measured by these tasks.

In healthy controls, the OSCARS yielded a 3-factor, rather than the 2-factor solution in the schizophrenia sample. This is consistent with findings on emotion intelligence, which showed a 4-factor model in healthy controls and a 2-factor model in people with schizophrenia (Eack, Pogue-Geile, Greeno, & Keshavan, 2009). This suggests that there might be qualitative differences in SC ability in controls and individuals with schizophrenia. The extent to which individuals with schizophrenia experience generalised versus specific SC deficits is not well understood, however, this might contribute to the present sample's differential factor analytic structures. Generalised deficits have been implicated in basic neurocognition, and likely result in a simpler factor structure (Dickinson & Harvey, 2009). Healthy controls may therefore have differentiated SC abilities, creating more variance, and hence, a greater number of factors.

The OSCARS showed preliminary evidence of external validity, as it was significantly, albeit modestly correlated with indices of functioning, particularly: global social functioning, working productivity, independent living, functionality total and approached statistical significance with QLS-Work. However, the OSCARS was not significantly correlated with role-play performance. Thus, the present data suggest that

the OSCARS is more consistently associated with critical functional abilities – the ability to perform basic self-care (e.g., cooking and cleaning), and to work and sustain employment, than abilities that manifest during social interactions. It is possible the OSCARS functions as more of a social capacity scale, whereby the scores indicate the level of SC an individual is capable of in an ideal situation, e.g., with an individual (informant) they see regularly and are comfortable with (Patterson & Mausbach, 2010).

The discriminant validity analyses showed that although the OSCARS association with IQ approached statistical significance, all OSCARS indices were significantly correlated with the observer-based SCoRS. Higher correlations between OSCARS and SCoRS may reflect that these measures are capturing similar constructs or that they are due to method variance, as the same informant provided information for both scales (discussed later).

Predictive validity analyses showed that the OSCARS Total score contributed unique variance to real-life functioning in individuals with schizophrenia, beyond that of the objective social cognitive measures. Overall, the OSCARS ratings predicted nearly twice the variance in functioning compared to objective measures of SC. The variance accounted for by SC in the present study is consistent with Fett et al.'s (2011) meta-analysis, where SC factor explained 16% of the variance in functioning, on average. The current findings indicate that OSCARS is providing unique supplementary information concerning patient level of functioning, beyond that of performance-based measures of SC.

Exploratory analyses revealed that first-degree family members rated individuals with schizophrenia as having higher levels of SC impairment than other informants. It is unclear if this difference between groups is as a result of error variance or true variance between groups. Potential error-related reasons for this difference may be related to: (1) first-degree family members' possible difficulty forming accurate ratings due to their own SC difficulties (Janssen, Krabbendam, Jolles, & Van Os, 2003) or (2) error related to frustration with the family member (Schulz et al., 2013). A true variance-related reason might reflect the first-degree family's increased frequency of contact with the individual, and thus they are better able speak to the individual's deficits. In NPCs, there were not significant differences between the two informant groups in OSCARS ratings.

The primary limitation of this study was that the same informant provided collateral information used to score the GSFS, RFS, SCoRS and OSCARS ratings. Thus, significant correlations may be partially due to common method variance, which measures systematic error. However, method variance does not account for near significant OSCARS relationships with non-observer-based scales SSPA 2 total, QLS-Work and WASI. Additionally, the RFS collects information on both social and non-social content (e.g., work and independent living), which decreases the likelihood that correlations are due to content similarity. Utilising different informants across observer-based measures would eliminate the possibility that method variance is responsible for significant relationships. Further, requiring different interviewers to gather collateral across informant-based scales would prevent potential contamination across scales of rating information. Thus, future work should explore the relationship between informant role and validity of OSCARS data.

In summary, this is the first known study to utilise informant report in the assessment of SC in individuals with schizophrenia. The OSCARS could provide supplemental collateral information beyond laboratory-based SC measures. OSCARS administration is brief (15–20 minutes) and appears to evidence external validity, though this may be due

to shared method variance. Further research is needed to better understand the OSCARS' relationships with real-world functioning. The present study provides preliminary evidence that the OSCARS may be useful for clinicians in collecting data about patients' real-world SC deficits, increasing the degree to which these impairments are considered treatment targets.

Acknowledgements

The authors thank Chris Wiesen, Ph.D. for his assistance in data analysis. The authors also thank Drs Piper Meyer, Sarah Uzenoff Mintz and Katy Harper Romeo, as well as Betty Rupp for their help in conducting this research.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the National Institute of Mental Health at the National Institutes of Health to David L. Penn [grant number 1-R34-MH080010-01].

References

Adolphs, R. (2009). The social brain: Neural basis of social knowledge. *Annual Review of Psychology*, *60*, 693–716. doi:10.1146/annurev.psych.60.110707.163514

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *38*, 813–822. doi:10.1111/j.1469-7610.1997.tb01599.x

Birchwood, M., Smith, J., Cochrane, R., Wetton, S., & Copestake, S. (1990). The Social Functioning Scale. The development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. *The British Journal of Psychiatry*, *157*, 853–859. doi:10.1192/bjp.157.6.853

Bora, E., Yucel, M., & Pantelis, C. (2009). Theory of mind impairment in schizophrenia: Meta-analysis. *Schizophrenia Research*, *109*(1–3), 1–9. doi:10.1016/j.schres.2008.12.020

Brekke, J., Kay, D. D., Lee, K. S., & Green, M. F. (2005). Biosocial pathways to functional outcome in schizophrenia. *Schizophrenia Research*, *80*, 213–225. doi:10.1016/j.schres.2005.07.008

Browne, M. W., Cudeck, R., Bollen, K. A., & Long, J. S. (1993). Alternative ways of assessing model fit. *SAGE Focus Editions*, *154*, 136.

Combs, D. R., Penn, D. L., Wicher, M., & Waldheter E. (2007). The Ambiguous Intentions Hostility Questionnaire (AIHQ): A new measure for evaluating hostile social-cognitive biases in paranoia. *Cognitive Neuropsychiatry*, *12*, 128–143. doi:10.1080/13546800600787854

Corcoran, R. (2003). Inductive reasoning and the understanding of intention in schizophrenia. *Cognitive Neuropsychiatry*, *8*, 223–235. doi:10.1080/13546800244000319

Cornblatt, B. A., Auther, A. M., Niendam, T., Smith, C. W., Zinberg, J., Bearden, C. E., & Cannon, T. D. (2007). Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. *Schizophrenia Bulletin*, *33*, 688–702. doi:10.1093/schbul/sbm029

Couture, S. M., Penn, D. L., & Roberts, D. L. (2006). The functional significance of social cognition in schizophrenia: A review. *Schizophrenia Bulletin*, *32*, S44–S63. doi:10.1093/schbul/sbl029

Dickinson, D., & Harvey, P. D. (2009). Systemic hypotheses for generalized cognitive deficits in schizophrenia: A new take on an old problem. *Schizophrenia Bulletin*, *35*, 403–414. doi:10.1093/schbul/sbn097

Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology*, *18*, 791–814. doi:10.1017/S0954579406060391

Dudley, R. E. J., John, C. H., Young, A. W., & Over, D. E. (1997a). The effect of self-referent material on the reasoning of people with delusions. *British Journal of Clinical Psychology*, *36*, 575–584. doi:10.1111/j.2044-8260.1997.tb01262.x

Dudley, R. E. J., John, C. H., Young, A. W., & Over, D. E. (1997b). Normal and abnormal reasoning in people with delusions. *British Journal of Clinical Psychology*, *36*, 243–258. doi:10.1111/j.2044-8260.1997.tb01410.x

Eack, S. M., Pogue-Geile, M. F., Greeno, C. G., & Keshavan, M. S. (2009). Evidence of factorial variance of the Mayer–Salovey–Caruso Emotional Intelligence Test across schizophrenia and normative samples. *Schizophrenia Research*, *114*(1–3), 105–109. doi:10.1016/j.schres.2009.05.011

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*, 272–299. doi:10.1037/1082-989X.4.3.272

Fett, A. J., Viechtbauer, W., Dominguez, M., 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*, 573–588. doi:10.1016/j.neubiorev.2010.07.001

Green, M. F., Olivier, B., Crawley, J. N., Penn, D. L., & Silverstein, S. (2005). Social cognition in schizophrenia: Recommendations from the measurement and treatment research to improve cognition in schizophrenia new approaches conference. *Schizophrenia Bulletin*, *31*, 882–887. doi:10.1093/schbul/sbi049

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*, 1211–1220. doi:10.1093/schbul/sbm145

Heinrichs, D. W., Hanlon, T. E., & Carpenter Jr, W. T. (1984). The Quality of Life Scale: An instrument for rating the schizophrenic deficit syndrome. *Schizophrenia Bulletin*, *10*, 388–398. doi:10.1093/schbul/10.3.388

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55.

Janssen, I., Krabbendam, L., Jolles, J., & Van Os, J. (2003). Alterations in theory of mind in patients with schizophrenia and non-psychotic relatives. *Acta Psychiatrica Scandinavica*, *108*(2), 110–117. doi:10.1034/j.1600-0447.2003.00092.x

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*, 261–276. doi:10.1093/schbul/13.2.261

Keefe, R. S. E., Poe, M., Walker, T. M., Kang, J. W., & Harvey, P. D. (2006). The Schizophrenia Cognition Rating Scale: An interview-based assessment and its relationship to cognition, real-world functioning, and functional capacity. *American Journal of Psychiatry*, *163*, 426–432. doi:10.1176/appi.ajp.163.3.426

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*, 312–318. doi:10.1037/0021-843X.102.2.312

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*, 1768–1774. doi:10.1176/appi.ajp.160.10.1768

Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., & Moberg, P. J. (2010). Facial emotion perception in schizophrenia: A meta-analytic review. *Schizophrenia Bulletin*, *36*, 1009–1019. doi:10.1093/schbul/sbn192

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. doi:10.1080/09638280600646185

McDonald, S., Flanagan, S., Rollins, J., & Kinch, J. (2003). TASIT: A new clinical tool for assessing social perception after traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, *18*, 219–238. doi:10.1097/00001199-200305000-00001

McPheeters, H. L. (1984). Statewide mental health outcome evaluation: A perspective of two southern states. *Community Mental Health Journal*, *20*(1), 44–55. doi:10.1007/BF00754103

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. doi:10.1016/j.schres.2004.09.007

Ochsner, K. N. (2008). The social-emotional processing stream: Five core constructs and their translational potential for schizophrenia and beyond. *Biological Psychiatry*, *64*(1), 48–61. doi:10.1016/j.biopsych.2008.04.024

Patterson, T. L., & Mausbach, B. T. (2010). Measurement of functional capacity: A new approach to understanding functional differences and real-world behavioral adaptation in those with mental illness. *Annual Review of Clinical Psychology*, *6*, 139–154. doi:10.1146/annurev.clinpsy.121208.131339

Patterson, T. L., Moscona, S., McKibbin, C. L., Davidson, K., & Jeste, D. V. (2001). Social skills performance assessment among older patients with schizophrenia. *Schizophrenia Research*, *48*, 351–360. doi:10.1016/S0920-9964(00)00109-2

Penn, D. L., Corrigan, P. W., Bentall, R. P., Racenstein, J. M., & Newman, L. (1997). Social cognition in schizophrenia. *Psychological Bulletin*, *121*(1), 114–132. doi:10.1037/0033-2909.121.1.114

Penn, D. L., Sanna, L. J., & Roberts, D. L. (2008). Social cognition in schizophrenia: An overview. *Schizophrenia Bulletin*, *34*, 408–411. doi:10.1093/schbul/sbn014

Pijnenborg, G. H. M., Withaar, F. K., Evans, J. J., Van den Bosch, R. J., Timmerman, M. E., & Brouwer W. H. (2009). The predictive value of measures of social cognition for community functioning in schizophrenia: Implications for neuropsychological assessment. *Journal of the International Neuropsychological Society*, *15*, 239–247. doi:10.1017/S1355617709090341

Pinkham, A. E., & Penn, D. L. (2006). Neurocognitive and social cognitive predictors of interpersonal skill in schizophrenia. *Psychiatry Research*, *143*, 167–178. doi:10.1016/j.psychres.2005.09.005

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*, 813–823. doi:10.1093/schbul/sbt081

Roberts, D., Combs, D. R., Willoughby, M., Mintz, J., Gibson, C. M., Rupp, B., & Penn, D. L. (2014). A randomized, controlled trial of Social Cognition and Interaction Training (SCIT) for outpatients with schizophrenia-spectrum disorders. *British Journal of Clinical Psychology*, *53*(3), 281–298.

Sabbag, S., Twamley, E. M., Vella, L., Heaton, R. K., Patterson, T. L., & Harvey, P. D. (2011). Assessing everyday functioning in schizophrenia: Not all informants seem equally informative. *Schizophrenia Research*, *131*, 250–255. doi:10.1016/j.schres.2011.05.003

Schulz, R., Cook, T. B., Beach, S. R., Lingler, J. H., Martire, L. M., Monin, J. K., & Czaja, S. J. (2013). Magnitude and causes of bias among family caregivers rating Alzheimer disease patients. *The American Journal of Geriatric Psychiatry*, *21*(1), 14–25. doi:10.1016/j.jagp.2012.10.002

Ventura, J., Cienfuegos, A., Boxer, O., & Bilder, R. (2008). Clinical global impression of cognition in schizophrenia (CGI-CogS): Reliability and validity of a co-primary measure of cognition. *Schizophrenia Research*, *106*(1), 59–69. doi:10.1016/j.schres.2007.07.025

Ventura, J., Reise, S. P., Keefe, R. S. E., Baade, L. E., Gold, J. M., Green, M. F., … Bilder, R. M. (2010). The Cognitive Assessment Interview (CAI): Development and validation of an empirically derived, brief interview-based measure of cognition. *Schizophrenia Research*, *121*(1–3), 24–31. doi:10.1016/j.schres.2010.04.016

Werner, P. D. (2001). Structured clinical interview for DSM-IV axis I disorders: Clinician version. In B. S. Plake & J. C. Impara (Eds.), *The fourteenth mental measurements yearbook* (pp. 1123–1125). Lincoln, NE: Buros institute of Mental Measurements.

Whitmyre, J. W., & Pishkin, V. (1958). The abbreviated Wechsler adult intelligence scale in a psychiatric population. *Journal of Clinical Psychology*, *14*, 189–191. doi:10.1002/1097-4679(195804)14:2<189::AID-JCLP2270140225>3.0.CO;2-4

Yager, J. A., Ehmann, T. S. (2006). Untangling social function and social cognition: A review of concepts and measurement. *Psychiatry: Interpersonal and Biological Processes*, *69*(1), 47–68. doi:10.1521/psyc.2006.69.1.47

Appendix

SCIT RCT STUDY 1

Subject ID: | | | | | | | **Visit #:** | | **Visit date: (MM/DD/YYYY)** | | | |

Subject Initials: | | |

OBSERVABLE SOCIAL COGNITION: A RATING SCALE (OSCARS)

David Penn, Ph.D., Dennis Combs, Ph.D., David Roberts, M.A., Steven Silverstein, Ph.D., and Richard Keefe, Ph.D.

Relationship with target person:_____

Interview Date:_____

PLEASE RATE WHETHER THE TARGET PERSON HAS DIFFICULTY IN EACH OF 8 AREAS USING THE SCALE PROVIDED BELOW. PLEASE RATE THE SEVERITY OF THEIR PROBLEMS FROM 1 (NONE) TO 7 (EXTREMELY SEVERE). THUS, HIGHER NUMBERS SUGGEST MORE SEVERE PROBLEMS.

RATE THE PERSON'S BEHAVIOUR OVER THE PAST 7 DAYS. CONSIDER ANY INTERACTIONS YOU HAVE HAD WITH HER/HIM AND ANY INTERACTIONS THAT YOU HAVE OBSERVED THAT HE/SHE HAS HAD WITH OTHER PEOPLE.

FOR EACH ITEM, WE PROVIDE A GENERAL EXAMPLE OF THE BEHAVIOUR IN QUESTION, THE RATING SCALE (1–7), THE LEVEL OF SEVERITY (NONE TO EXTREMELY SEVERE) AND BEHAVIORAL DESCRIPTORS THAT CORRESOND TO THE RATINGS. PLEASE RATE ALL ITEMS. INDICATE YOUR RATINGS BY CIRCLING THE NUMBER ON THE SCALE FOR THAT PARTICULAR ITEM. YOU CAN USE ANY NUMBER ON THE SCALE.

LEVEL OF SEVERITY RATING SCALE

1	2	3	4	5	6	7
None	Very Mild	Mild	Moderate	Moderately Severe	Severe	Extremely Severe

SCIT RCT STUDY 1

Subject ID: | | | | | | | **Visit #:** | | **Visit date: (MM/DD/YYYY)** | | | |

Subject Initials: | | |

Does the individual have difficulty in the following areas?

1. Recognising other people's emotions, particularly negative emotions (sadness, fear and anger) based on facial expression, body language and/or vocal tone and rate?

EXAMPLE: When talking to someone, the individual cannot tell that the other person is upset or angry. They seem "clueless" about how other people are feeling.

1	**None**	Can recognise strong, moderate and subtle expressions of emotions. She/he can be thought of as "socially perceptive".	
2	**Very Mild**		**Informant**
3	**Mild**	Easily recognises strong and explicit expressions of emotion, such as crying, angry shouting or elated laughing. May not recognise moderately expressed emotions. However, she/he does not recognise subtle expressions, such as disappointment expressed only with a sigh or slight mouth movement.	**1 2 3 4 5 6 7**
4	**Moderate**		**Interviewer**
5	**Moderately Severe**	Recognises most strong and explicit expressions. Does not recognise moderate or subtle emotional expressions	
6	**Severe**		**1 2 3 4 5 6 7**
7	**Extremely Severe**	Never or does not recognise strong, moderate and subtle emotional expressions. The person must be told what emotion is being expressed ("I am very angry.")	

2. Interpreting social interactions in a malevolent, hostile manner?

EXAMPLE: The individual sees others as intending them harm, especially in ambiguous (unclear) situations. For example, she/he walks past a few people who are laughing and thinks that they are laughing at her/him. Or, the individual can't find a personal item and thinks that someone else took it. (Ask about frequency)

1	**None**	Individual does not immediately blame others or think that they have ill intentions towards them. She/he will consider the possibility that other factors, such as situational factors, may have caused the outcome (e.g., that others are laughing at a joke, not them).	
2	**Very Mild**		**Informant**
3	**Mild**	Sometimes feels that others intend them harm that may be transient or short-lived	**1 2 3 4 5 6 7**
4	**Moderate**		
5	**Moderately Severe**	Frequently interprets others behaviour as ill intentioned. May sometimes accuse close acquaintances of ill will.	**Interviewer**
6	**Severe**		**1 2 3 4 5 6 7**
7	**Extremely Severe**	Widespread beliefs of ill will, which includes both close acquaintances and strangers. These beliefs are generally unfounded.	

SCIT RCT STUDY 1

Subject ID: | | | | | | | Visit #: | | Visit date: (MM/DD/YYYY) | | | |

Subject Initials: | | |

3. Making decisions quickly (i.e., jumps to conclusions) without examining other evidence?

EXAMPLE: The individual attempts to call you and you do not answer. They immediately believe that you did not pick up because you are mad at them. (Ask about frequency)

1	**None**	Does not rush to decisions; is thoughtful and deliberate. The person seeks out other information and takes time to carefully weigh the pros and cons before making a decision.	
2	**Very Mild**		**Informant**
3	**Mild**	Sometimes uses only immediate information to make decisions. The person makes decisions using additional information some of the time and can weigh the pros and cons if motivated.	**1 2 3 4 5 6 7**
4	**Moderate**		
5	**Moderately Severe**	Often uses only immediate information to make decisions. The person must be prompted or told to examine other sources of evidence and take their time in making decisions.	**Interviewer**
6	**Severe**		**1 2 3 4 5 6 7**
7	**Extremely Severe**	Frequently uses only the most immediate information to make decisions. The person does not seek out additional information and seems to rush to judgement almost without thinking. Seems irrational.	

4. Being flexible in interpreting social situations?

EXAMPLE: The individual is waiting for someone at a restaurant and they are 20 minutes late. They conclude that the person does not want meet them when in reality they may be stuck in traffic or have had a flat tire. They do not consider other alternatives for why the person is late and stick to one opinion.

1	**None**	When considering someone else's behaviour, she/he is able to come up with multiple reasons for why she/he acted the way they did.	
2	**Very Mild**		**Informant**
3	**Mild**	Has some trouble coming up with guesses, but can do so if motivated.	**1 2 3 4 5 6 7**
4	**Moderate**		
5	**Moderately Severe**	Has difficulty even thinking up multiple possible explanations for others' social behaviour; Has to be prompted or asked to come up with different guesses about another person's behaviour.	**Interviewer**
6	**Severe**		**1 2 3 4 5 6 7**
7	**Extremely Severe**	Does not understand that more than one interpretation of an event is possible and is unable to generate any alternative guesses.	

SCIT RCT STUDY 1

Subject ID: | | | | | | | **Visit #:** | | **Visit date: (MM/DD/YYYY)** | | | |

Subject Initials: | | |

5. Can change or correct their interpretation of social interactions when wrong?

EXAMPLE: The individual sees two people whispering and they believe they are talking about them. This belief is maintained even when told that one of the people was sharing something personal about themselves.

1	**None**	Is able to seek out, and weigh, evidence for and against a given belief about someone else.	
2	**Very Mild**		**Informant**
3	**Mild**	The individual will consider evidence that contradicts misinterpretations she or he has made, although they might maintain the false belief anyway.	**1 2 3 4 5 6 7**
4	**Moderate**		
5	**Moderately Severe**	The individual often avoids listening to facts that contradicts his/her views, or may argue strongly against them, and usually maintains the false belief.	**Interviewer**
6	**Severe**		**1 2 3 4 5 6 7**
7	**Extremely Severe**	The individual refuses to consider contradictory evidence. It feels impossible to talk the individual out of a belief even when the belief is clearly wrong.	

6. Understanding subtle jokes, sarcasm and insults in conversation?

EXAMPLE: Someone states during a meal that this is best food I have ever had in a sarcastic tone and the individual does not realise that this is an insult and/or sarcasm.

1	**None**	The individual understands subtle jokes, insults or sarcasm.	
2	**Very Mild**		**Informant**
3	**Mild**	The individual sometimes doesn't get subtle jokes or insults made by others (e.g., "Oh yes, I *love* working 15 hours a day!"). Seems to take longer to get the meaning of jokes and sarcasm.	**1 2 3 4 5 6 7**
4	**Moderate**		
5	**Moderately Severe**	The individual often does not understand subtle jokes, sarcasm, or insults and must be told what they mean.	**Interviewer**
6	**Severe**		**1 2 3 4 5 6 7**
7	**Extremely Severe**	The individual does not understand subtleties at all. Statements must be concrete and direct in order to be understood (e.g., slapstick humour).	

SCIT RCT STUDY 1

Subject ID: | | | | | | | Visit #: | | Visit date: (MM/DD/YYYY) | | | |

Subject Initials: | | |

7. Seeing things from the perspective of others (i.e., putting themselves in other people's shoes)?

EXAMPLE: The individual cannot understand why someone feels upset or angry in a particular situation. Or, when watching a sad film, does not feel moved by it.

1	**None**	Individual can be described as empathic. When watching a sad or happy film (or reading a sad or happy book), can be moved by it.	
2	**Very Mild**		**Informant**
3	**Mild**	Seldom inquires or makes guesses about others mental states or feelings (e.g., "Do you like that?" or "Are you confused?"), but understands when people make these statements.	**1 2 3 4 5 6 7**
4	**Moderate**		
5	**Moderately Severe**	Only reacts empathically when others express strong emotion. Does not ask about or respond to others' opinions or experiences. When asked, may have difficulty imagining what others might be thinking.	**Interviewer**
6	**Severe**		**1 2 3 4 5 6 7**
7	**Extremely Severe**	Unable to accurately judge what others might be thinking or feeling, except in the most extreme circumstances (such as feeling sadness after the death of a loved one.)	

8. Understanding subtle social cues, hints and indirect requests (an example of an indirect request is if your son/daughter wants a toy, but rather than say so directly, comments on how pretty it is.

EXAMPLE: You are trying to read a book or watch TV and the individual keeps talking to you, even though you are giving off subtle hints/signals that you are not interested in talking to them at that moment (e.g., keeping your answers short; not making eye contact).

1	**None**	The individual readily picks up social cues and/or indirect requests. For example, if you are busy and they start talking to you, they readily perceive that you can't speak with them at that moment.	
2	**Very Mild**		**Informant**
3	**Mild**	The individual does not pick up on subtle social cues at first, but does so after a minute or two. Takes longer to process subtle cues and hints.	**1 2 3 4 5 6 7**
4	**Moderate**		
5	**Moderately Severe**	The individual does not pick up on social cues and it takes a number of overt cues (turning away when talking to him/her) for them to get the message.	**Interviewer**
6	**Severe**		**1 2 3 4 5 6 7**
7	**Extremely Severe**	The individual does not pick up on social cues and must be told directly. Or, the individual does not get subtle hints or indirect requests.	